Wave-equation Traveltime-Offset Inversion Using a Bi-dimensional Cross-correlation Function

Olga Yurancy Savasta Oropeza

Doctor of Philosophy

The University of Edinburgh

1998

To:

Victor Varela V. (my husband)

Victor Varela G. (his father)

My brothers and sisters

Abstract

This dissertation addresses the problem of inverting two-dimensional seismic data to determine the compressional wave velocity as a function of the spatial position in the medium. An automatic procedure for velocity estimation is set up using a bi-dimensional cross-correlation function to compare the modelled seismograms with the data. The standard wave-equation traveltime inversion algorithm treats each recorded seismic trace independently, and ignores the offset-dependence of reflected and diffracted arrivals in an ensemble of traces in a shot gather. My new inversion method, wave-equation traveltime-*offset* inversion, focuses on reflected or diffracted waveforms and introduces the bi-dimensional (offset-time) cross-correlation function to tackle this offset dependence. The velocity model is updated using the current velocity model, the observed pressure field, and the fields computed by reverse time propagation of two pseudo-residual functions acting as sources in a particular location.

The wave-equation traveltime-offset inversion reconstructs the low frequency content of the velocity model when transmitted arrivals are used. Combined with the full waveform inversion, it succeeds in inverting a synthetic fault model in the crosshole configuration. The proposed method also succeeds in inverting surface reflection datasets while the standard traveltime inversion fails. By taking into account the moveout between traces, the convergence is more stable than with the conventional method. When inverting datasets with surface reflection geometries, the dominant event and the velocity above it are recovered using the traveltime and offset information. The remaining interfaces are defined as well. However, as there is a gap in resolution between the long and short wavelengths, the blocky variations in the velocity model are not observed.

Acknowledgements

Starting from the beginning, I have to thank Orlando Chacín for proposing my name to the scholarship committee and for believing I could make it. I thank Intevep, S.A for sponsoring my PhD studies and Carmen Murgich for the administrative support. I am very grateful to my supervisors Prof. Anton Ziolkowski and Dr. Robert Pearce, for their support, advice and encouragement; they also checked this dissertation very carefully. I cannot find the right word to express my gratitude to my husband, Victor Varela, for his encouragement, his support, his love, and for the continuous suggestions and comments he gave me in order to improve the results of this work.

I thank Prof. G. Schuster and Dr. C. Zhou for providing the full-wave inversion computer code and the crosshole velocity model. I am grateful to Dr. A. Pica and Dr. S. Zatsepin for helpful discussions. The help of Prof. A. Tarantola and the people in the Equipe de Tomographie Geophysique, Institut de Physique du Globe de Paris is also acknowledged; special thanks go to Miguel Bosch. Andrey Ortega helped me from Dallas University. I also thank the Edinburgh Parallel Computing Centre group (EPCC) for their support.

I specially thank Mrs. Margaret Penman, for being so kind and for her invaluable help. My gratitude goes to my English teacher, Sue Tait, for being so patient while trying to show me the difference in the pronunciation of "cup" and "cap", and "ankle" and "uncle" (among many others words). Sue also checked the grammar and the spelling of this dissertation, together with Chung-Lun Lau and Nick Braun. Chung also helped me a lot with the computing problems and Steve Law was in charge of my computer at home.

I am very grateful to my sister Maritza for all her support and help in the administration of my things in Venezuela, and to Debora Cores for her support and for being the contact between Maritza and Intevep. Also, the encouragement from my sisters, my brothers, my nieces and nephews is acknowledged.

I appreciate very much the help and support of Sara Buchanan, Mark Higgins,

Rodney Johnston, Orlando Oliveira, Francisco Vieira and Andrew Wilson. I have to mention the Women's Club where I met wonderful people from all over the world and where I also met most of my best friends in Edinburgh; special thanks go to Mrs. Bo Upton and Mrs. Wendy Rutter. Finally, my thanks go to all my friends (a very long list) for their support, encouragement and for making my life as a foreigner much more pleasant.

Contents

List of Figures

Conventions, notation, definitions and symbols

Notations

CRP	common reflection point
Cray T3D	parallel computer
EPCC	Edinburgh Parallel Computing Centre
Maple	mathematical calculation package (Maple V Release 4)
MPI	message passing interface
ProMAX	2D seismic processing package (version 6.0)
2D	two-dimensional
3D	tri-dimensional

Conventions

$\tilde{}$	complex conjugate
$*$	convolution
$\otimes$	cross-correlation
Δ	element
∇	gradient
$\nabla\cdot$	divergence
$\hat{}$	dual space
$\mathbf{C^t}$	transpose operator
$\mathbf{C^{-1}}$	inverse operator
F	Fourier transform of f
$x = (x, y, z)$	tri-dimensional position vector
$x = (x, z)$	two-dimensional position vector
$x_r = (x, 0)$	receiver position vector for surface reflection geometry configuration
$x_r = (x, z)$	receiver position vector for crosshole geometry configuration
$x_s = (x, 0)$	source position vector for surface reflection geometry configuration
$x_s = (x, z)$	source position vector for crosshole geometry configuration

Conventions ....

$\dot{f} = \frac{\partial f}{\partial t}$	derivative of f with respect to time
$\ddot{f} = \frac{\partial^2 f}{\partial t^2}$	second derivative of f with respect to time
$f' = \frac{\partial f}{\partial x}$	derivative of f with respect to the receiver position coordinate
$f'' = \frac{\partial^2 f}{\partial x^2}$	second derivative of f with respect to the receiver position coordinate
$\dot{f}' = \frac{\partial^2 f}{\partial t \partial x}$	mixed derivative of f with respect to time and the receiver position coordinate
$\int dt\, f = \int_{-\infty}^{\infty} dt\, f = \int_0^T dt\, f$	integral of $f(t)$ over the time, but as the seismic traces have finite length the limits run from 0 to the maximum recorded time T.
$\int dx\, f = \int_{-\infty}^{\infty} dx\, f = \sum_{x_1}^{x_2} f$	integral of $f(x)$ over the receiver position coordinate, but the seismic shot gathers have finite offset, running from the receiver in the position x_1 up to the receiver in the position x_2

Definitions

$f(t) \otimes g(t) = \int_{-\infty}^{\infty} dt\, f(t+\tau) g(t)$	one-dimensional cross-correlation function
$f(x,t) \otimes g(x,t) = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dx\, f(x+\sigma, t+\tau) g(x,t)$	bi-dimensional cross-correlation function
$F(x,\omega) = \int_{-\infty}^{\infty} dt f(x,t) e^{-i\omega t}$	time-Fourier transform
$f(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega F(x,\omega) e^{i\omega t}$	inverse time-Fourier transform
$\bar{F}(k,\omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dt\, dx f(x,t) e^{ikx} e^{-i\omega t}$	space-time Fourier transform
$f(x,t) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\omega\, dk \bar{F}(k,\omega) e^{-ikx} e^{i\omega t}$	inverse space-time Fourier transform
$s(t; x_s) = s(t)\delta(x - x_s)$	source function

Definitions ....

δ^{ab} Kronecker's delta (1 if $a = b$; 0 otherwise)

Symbols

$A(x_r; x_s)$	$[Nm^{-2}]$	maximum amplitude of each individual trace
$\mathbf{C_p}$		covariant operator
$\mathbf{c}$		model space
$c(x)$	$[(m\,s^{-1}]$	compressional wave velocity
$F = (F_x, F_y, F_z)$	$[N]$	force vector
$G(x,t)$	$[kg\,m^{-3}s^{-1}]$	rate of mass change per unit volume
$\mathbf{G_c}$		Fréchet derivative operator
$g(x,t;x',t')$	$[m^{-1}s^{-1}]$	Green's function
$\mathbf{I}$		identity operator
K	$[kg\,m^{-1}\,s^{-2}]$	compression or bulk modulus
$k_r = (k_x, k_z)$	$[m^{-1}]$	horizontal wavenumber
k_z	$[m^{-1}]$	vertical wavenumber
m	$[kg]$	mass
nx		number of traces in a shot gather
nt		number of samples in time
$\mathbf{p}$		seismic dataset space
$p(x,t;x_s)$	$[Nm^{-2}]$	pressure field in the position x, and generated by a source in the position x_s
$p(x,t;x_s)_{obs}$	$[Nm^{-2}]$	observed seismogram
$p(x,t;x_s)_{cal}$	$[Nm^{-2}]$	calculated seismogram
$Q(x,t;x_s)$	$[kg^{-1}m\,s]$	back-propagated field generated by the offset pseudo-residual acting as a source in the position x_r
$q_1(x,t;x_s)$	$[kg^{-1}s^3]$	back-propagated field generated by the standard traveltime residual acting as a source in the position x_r

Symbols ....

$q_2(x,t;x_s)$	$[kg^{-1}s^2]$	back-propagated field generated by the seismogram residual acting as a source in the position x_r
$R(x,t;x_s)$	$[kg^{-1}m^{-1}s^3]$	back-propagated field generated by the traveltime pseudo-residual acting as a source in the position x_r (traveltime-offset inversion)
S_x	$[m^2]$	surface perpendicular to x axis
S_y	$[m^2]$	surface perpendicular to y axis
S_z	$[m^2]$	surface perpendicular to z axis
S_1	$[s^2]$	misfit function for the standard traveltime inversion
S_2		weighted misfit function for the full waveform inversion
S	$[m^2]$	misfit function for the traveltime-offset inversion
$\mathbf{S_0}$		preconditioning factor operator
T	$[s]$	recorded time
t	$[s]$	time coordinate
$u=(u_x,u_y,u_z)$	$[m\,s^{-1}]$	particle velocity vector
$u_s(x)$	$[m^{-1}s]$	slowness
$u_{s_0}(x)$	$[m^{-1}s]$	reference slowness
v	$[m^3]$	volume
X	$[m]$	maximum offset
α		step length
β	$[m^2s^{-2}]$	parameter to keep the dimension consistent in the misfit function S
Δt	$[s]$	time step
$\Delta\tau$	$[s]$	time delay where the cross-correlation function has a maximum

Symbols ....

$\Delta\tau_{max}$	$[s]$	maximum time delay allowed during the inversion
Δx	$[m]$	grid size
$\Delta\sigma$	$[m]$	offset delay where the cross-correlation function has a maximum
$\Delta\sigma_{max}$	$[m]$	maximum offset shift allowed during the inversion
$\delta(x - x_s)$		Dirac's delta function centred in x_s
$\delta\mathbf{p}$		difference between the current and the observed pressure field in the seismic data space
$\delta p(x_r, t; x_s)$	$[Nm^{-2}]$	seismogram residual
$\delta\hat{p}(x_r, t; x_s)$	$[N^{-1}m^2]$	seismogram residual in the dual space
$\delta\tau(x_r, t; x_s)$	$[kg^{-1}m\,s^3]$	standard pseudo-traveltime residual
$\delta\tau(x_r, t; x_s)$	$[kg^{-1}s^3]$	pseudo-traveltime residual in the traveltime-offset inversion
$\delta\sigma(x_r, t; x_s)$	$[kg^{-1}m^2s]$	pseudo-offset residual
$\gamma_1(x)$	$[m^{-4}s^3]$	steepest descent direction of the misfit function S_1
$\gamma_2(x)$		steepest descent direction of the misfit function S_2
$\hat{\gamma}_2(x)$	$[m^{-4}s^2]$	steepest descent gradient of the misfit function S_2
$\gamma(x)$	$[m^{-2}s]$	steepest descent direction of the misfit function S
ν	$[Hz]$	frequency
$\Phi(x)$	$[m]$	preconditioning function
ρ	$[kg\,m^{-3}]$	density
σ	$[m]$	offset shift

Symbols ....

σ_{sd}^2	$[N^2m^{-4}]$	standard deviation
τ	$[s]$	time delay
ω	$[rad\,s^{-1}]$	angular frequency

Chapter 1

Introduction

1.1 The problem

The characterisation of hydrocarbon reservoirs depends upon the imaging of the geological structures beneath the surface. Currently the process of imaging complex structures is too slow while decisions for the development of new reservoirs need to be taken quickly. A rapid and precise velocity estimation would allow reservoirs to be identified in a short time and in an accurate way, thus reducing cost.

The best imaging is achieved when the correct velocity model is used for the migration of the seismic data. A distorted structure results if the wavefield is propagated with erroneous velocities; even minor velocity errors can severely affect the migrated data. To overcome this problem several methods for velocity estimation have emerged in the last years, but most of them take too long to compute a reasonable velocity field, and human intervention is required.

The principal aim of this dissertation is to generate an algorithm for automatic velocity estimation, enabling the correct imaging of the structures to be done quickly, and keeping the human intervention to a minimum. This would allow a rapid evaluation of the production potential of the hydrocarbon reservoirs. Specifically, this dissertation addresses the problem of inverting 2D seismic data to determine the compressional wave velocity as a function of the spatial position in the medium.

1.2 Previous work on velocity estimation

A variety of methods for non-conventional velocity analysis has appeared in the last 15 years. At least four different domains have been used to search for the velocity model that best images the earth structure. Among them, I can mention:

- Focusing depth error panels.
- Coherency panels and migrated CRP gathers.
- Post-migrated depth or image domain.
- Pre-migrated time or data domain.

Depth focusing analysis is an iterative procedure that updates the velocity model using focusing errors (MacKay and Abma, 1992; Jeannot and Berranger, 1994). When the migration velocities are erroneous the imaging conditions are not fulfilled, and the energy is focused at non-zero times. Zero-offset data at non-zero times are preserved during migration, forming a large volume of data from which a depth-error surface can be extracted. Once the depth error is known, the velocity field can be updated. The beam-stack method is another iterative approach that maximises the coherent energy along hyperbolic trajectories. The coherent energy is transformed into semblance peaks that are used to find the velocity model that best predicts the corresponding semblance peaks. The resolution is restricted to the low wavenumbers of the velocity function, and it is constrained by the ray tracing coverage.

The horizontal alignments or non-alignments of the offset images depend on the errors in the velocity field. Therefore, coherency panels in the image domain (migrated common reflection point gathers) can be used to validate the velocity model. Some approaches maximise a scalar similarity index associated with a set of migrated sections (Symes and Carazzone, 1991; Chavent and Jacewitz, 1995). A certain normalised stack criterion discriminates errors in the velocity models

using the sum of the energies of the migrated shot records over all shot locations (Versteeg and Marfurt, 1994).

In the image domain, the difference of the imaged depths (residual moveout) at a common image gather has been used to update the velocity field (Liu and Bleistein, 1995). Also, the difference between depth-migrated shot gathers of pairs of neighbouring shot gathers in a least-squares sense has been implemented as a migration misfit criterion (Varela et al., 1994). Moreover, a method of estimation of 3D velocity models using 3D post-stack depth migration in an iterative, quasi-layer stripping method has been proposed (Jones, 1994).

This dissertation primarily focuses on the data domain, in which several works have been done in term of: traveltime and diffraction tomography inversions (Dines and Lytle, 1979; Bishop et al., 1985; Ivansson, 1985; Leggett et al., 1993; Varela et al., 1994), full waveform inversion (Tarantola, 1984; Tarantola, 1986; Gauthier et al., 1986; Mora, 1987a; Mora, 1987b; Mora, 1988; Mora, 1989; Jannane et al, 1989; Pica et al., 1990; Pratt and Goulty, 1991; Xu et al., 1995, among others), and wave-equation traveltime inversion (Luo and Schuster, 1991; Zhou et al., 1995). Also, tomographic methods based on the local coherency operators in the data domain have been proposed: beam-stack method (Biondi, 1992), and coherency inversion (Landa et al., 1991).

The misfit function to be minimised in traveltime tomography is defined as the sum of the squared errors between the observed and the calculated traveltimes. In this case, traveltime picks must be provided. In addition, high frequency approximations are implicit in the method. This procedure is usually applied on crosswell datasets involving direct or transmitted arrivals. However, some attempts to include reflection arrivals for velocity inversion together with the reflector position inversion have been made (Bishop et al., 1985; Stork and Clayton, 1986).

Full waveform inversion uses the normed difference between the observed and the calculated seismograms, and should be applied when the initial model is close to the actual model. Even though this method does not have high frequency res-

trictions, the misfit function can be highly non-linear with respect to the velocity perturbations.

The wave-equation traveltime inversion overcomes the restrictions imposed on the methods described above. No traveltime picks are necessary and no approximations are imposed on the data. This method has been applied to reconstruct the long wavelength features of the velocity field using datasets with crosshole configuration. Therefore, it has to be combined with the full waveform inversion to recover the fine details present in the velocity model.

1.3 The specific problem and the outline of this dissertation

The critical step for automating the process of finding the velocity model is the determination of the error criteria. Usually, errors are minimised using an optimisation algorithm, in such a way that the best velocity model that reproduces the real data is estimated. A good choice, in the data domain, is the traveltime delay that maximises the cross-correlation function between the real and the calculated data.

The traveltime inversion is based on this criterion, and it has been successfully applied to crosshole datasets muting out all waveforms except those of the transmitted arrivals. In this dissertation, I apply the same procedure to reflection and diffracted waveforms muting out the first arrivals in surface reflection datasets. I show that artifacts are present in the resulting velocity model, and the convergence is unstable. The standard traveltime inversion uses the cross-correlation function to evaluate the degree to which the calculated and the observed traces match each other. As the traces in a shot gather are shifted independently, the offset-dependence of the reflected or diffracted waveforms is ignored.

I propose a new inversion method, the wave-equation traveltime-offset inversion, to tackle the mentioned offset-dependence. I introduce a bi-dimensional cross-correlation function to estimate the similarity index between the calculated

and the observed shot gathers in the time and offset domain simultaneously. The new misfit criterion includes an offset-dependence term. I present the traveltime-offset inversion as a new approach to estimate a velocity model that generates calculated shot gathers that best match the observed dataset. I show the results of applying the new method to synthetic datasets.

As the traveltime-based methods recover only the long wavelength features present in the velocity model, the full waveform inversion is introduced to reconstruct the fine details of the velocity field. The hybrid methods (traveltime-full waveform, and traveltime-offset-full waveform) succeed in the inversion of a crosshole dataset. However, I show how some parameters can severely affect the results when inverting synthetic datasets with surface reflection geometry.

The outline of this dissertation is as follows. Chapter 2 includes a review of the standard traveltime inversion, and shows the results of inverting datasets with crosshole and surface reflection geometries, respectively. In Chapter 3, a description of the full waveform inversion is given, and the effect of some parameters on the inversion is analysed. Chapter 4 is devoted to the derivation of the equations that build up the new traveltime-offset inversion method. An overview of the different components involved in the numerical simulation is given in Chapter 5. Chapter 6 contains examples of the application of the new method. The results are discussed, emphasising the strengths and the weaknesses of the new traveltime-offset inversion. Chapter 7 summarises the main results of this dissertation, and suggestions for future work are given.

As my initial scheme for the automated velocity estimation included a migration algorithm (see Figure 1.1) I started working on the implementation and parallelisation of the split-step Fourier migration and the Fourier finite-difference migration codes in 1995. These algorithms were tested on synthetic datasets. Simultaneously, Tanis and Stoffa were working on the parallelisation of the split-step Fourier migration. They published their results in 1997. I enclose my report as an appendix in this dissertation.

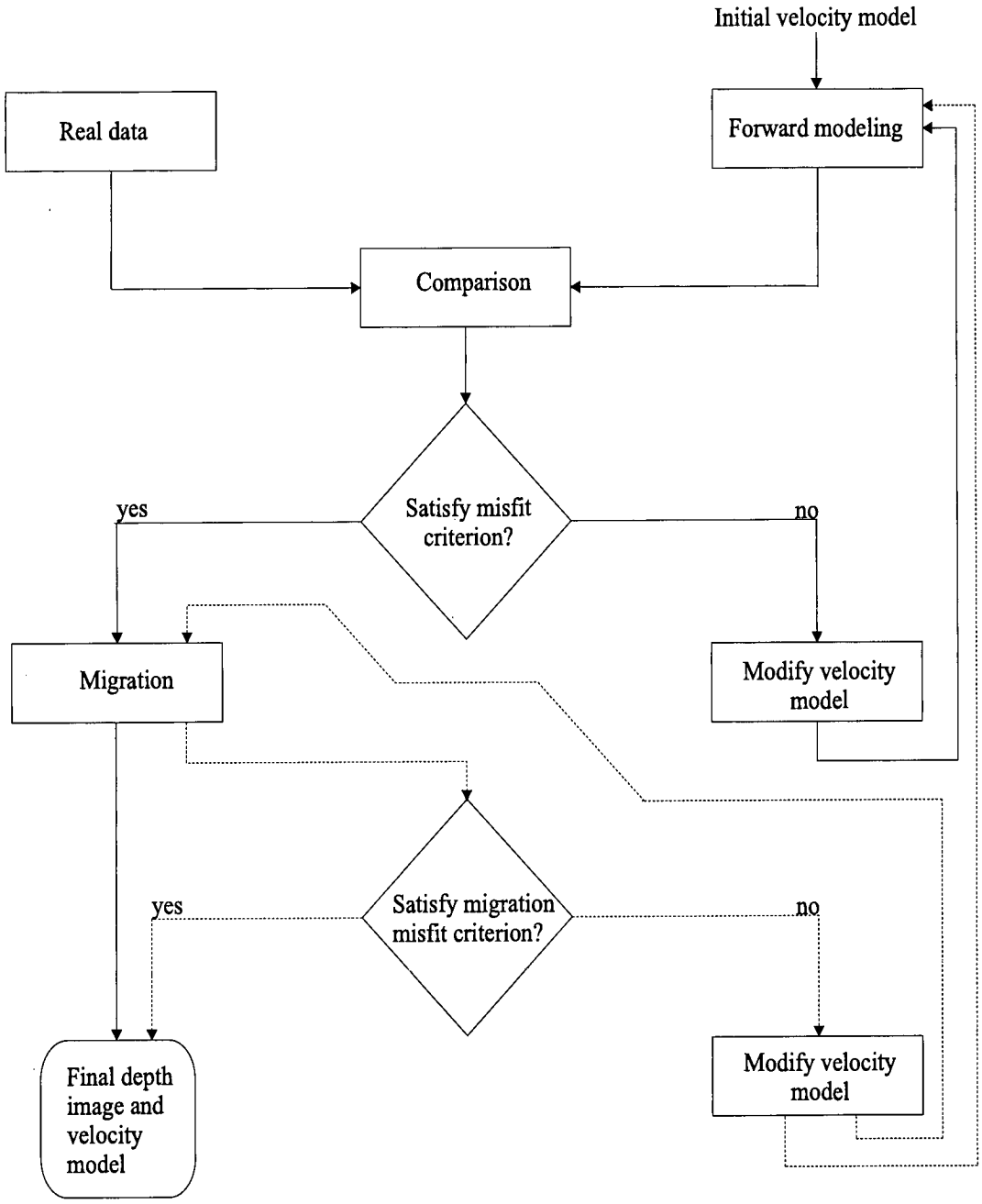


Figure 1.1: General scheme for the automated velocity analysis.

Chapter 2

Standard wave-equation traveltime inversion

Among the wave-equation inversion methods, the traveltime inversion has been used to reconstruct the low frequency content of the velocity model. It has the advantage that no high-frequency approximations are made, and no traveltime picking is necessary. Also, it can consider the starting models to be moderately far from the real model. Some weaknesses associated with this method are overcome by combining it with the full waveform inversion that provides the short wavelength features of the model.

Accurate results are obtained for crosshole configurations when transmitted arrivals are inverted. However, the inversion of surface reflection datasets shows that the standard traveltime inversion is not efficient enough and it can diverge if an appropriate step length is not used.

2.1 The acoustic wave equation

The wave-equation inversion methods are based on the acoustic wave equation. The derivation of the equation assumes the propagation of compressional waves in an inhomogeneous, isotropic fluid with zero viscosity and no heat conduction (Claerbout, 1976; Kinsler et al., 1982; Berkhout, 1985).

Considering a volume element Δv of mass Δm within the fluid as depicted in Figure 2.1, the pressure variations at opposite sides of the volume will accelerate it according to Newton's second law

$$\Delta F(x) = \Delta m \frac{\partial u(x,t)}{\partial t}, \tag{2.1}$$

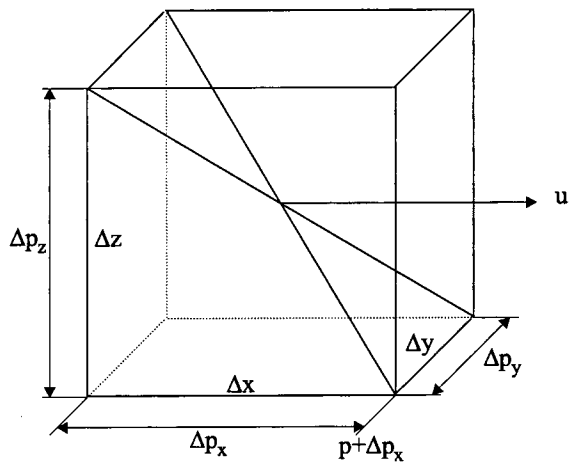


Figure 2.1: The average velocity within volume Δv of mass Δm will change due to spatial pressure variations (taken from Berkhout, 1985).

with $u(x,t)$ the particle velocity, and the components of the force ΔF satisfying

$$\begin{aligned}
\Delta F_x &= -\Delta p_x S_x = -\left(\frac{\partial p(x,t;x_s)}{\partial x}\Delta x\right)\Delta y\Delta z,\\
\Delta F_y &= -\Delta p_y S_y = -\left(\frac{\partial p(x,t;x_s)}{\partial y}\Delta y\right)\Delta x\Delta z,\\
\Delta F_z &= -\Delta p_z S_z = -\left(\frac{\partial p(x,t;x_s)}{\partial z}\Delta z\right)\Delta x\Delta y.
\end{aligned} \tag{2.2}$$

Hence, equation (2.1) can be written as

$$-\frac{1}{\rho(x)}\nabla p(x,t;x_s) = \frac{\partial u(x,t)}{\partial t}, \tag{2.3}$$

where $\rho = \Delta m/\Delta v$ is the density of the fluid.

Consider now that the mass Δm is exposed to an external force, then its volume Δv will change (see Figure 2.2). Hooke's law establishes the relationship between the pressure and the volume changes

$$dp(x,t;x_s) = -K(x)\frac{dv}{\Delta v}, \tag{2.4}$$

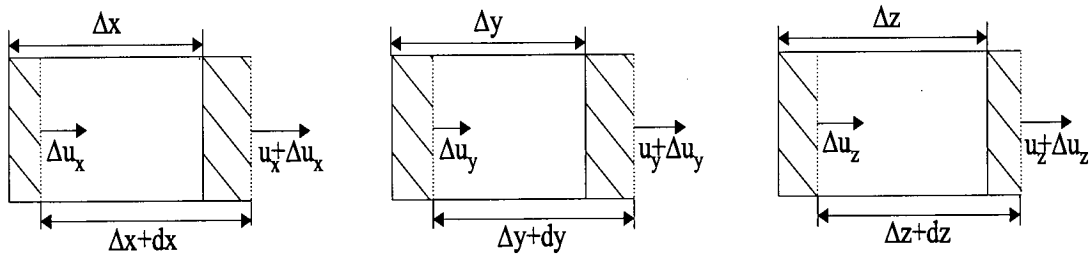


Figure 2.2: The average pressure within volume Δv of mass Δm will change due to spatial velocity variations (taken from Berkhout, 1985).

where $K(x)$ represents the compression modulus and $dv/\Delta v$ is given by

$$\frac{dv}{\Delta v} = \frac{dx}{\Delta x} + \frac{dy}{\Delta y} + \frac{dz}{\Delta z}, \tag{2.5}$$

with

$$\begin{aligned} dx &= \frac{\partial}{\partial x}(u_x dt)\Delta x, \\ dy &= \frac{\partial}{\partial y}(u_y dt)\Delta y, \\ dz &= \frac{\partial}{\partial z}(u_z dt)\Delta z, . \end{aligned} \tag{2.6}$$

Sustituting equations (2.5) and (2.6) into (2.4) results in

$$dp(x,t;x_s) = -K(x)\nabla \cdot u(x,t)\, dt, \tag{2.7}$$

and using $dp = \frac{\partial p}{\partial t} dt$, gives

$$\frac{\partial p(x,t;x_s)}{\partial t} = -K(x)\nabla \cdot u(x,t). \tag{2.8}$$

Applying the divergence operator to equation (2.3)

$$\nabla \cdot \left[-\frac{1}{\rho(x)} \nabla p(x,t;x_s) \right] = \frac{\partial}{\partial t} \nabla \cdot u(x,t), \tag{2.9}$$

and introducing equation (2.8) results in the acoustic wave equation

$$\nabla \cdot \left[\frac{1}{\rho(x)}\nabla p(x,t;x_s)\right] = \frac{1}{K(x)}\frac{\partial^2 p(x,t;x_s)}{\partial t^2}. \tag{2.10}$$

The propagation velocity of the compressional waves $c(x)$ is written as a function of the compression or bulk modulus as

$$c(x) = \sqrt{\frac{K(x)}{\rho(x)}}, \tag{2.11}$$

hence, equation (2.10) reduces to the usual form of the homogeneous wave equation

$$\frac{1}{c(x)^2}\frac{\partial^2 p(x,t;x_s)}{\partial t^2} - \rho(x)\nabla \cdot \left[\frac{1}{\rho(x)}\nabla p(x,t;x_s)\right] = 0. \tag{2.12}$$

The previous equation does not take into account external sources. However, a source function is needed to generate acoustic disturbances (Kinsler et al., 1982; Sheriff and Geldart, 1995). There are different types of sources, but let us consider here only two of them:

a) if body forces are considered, then this force per unit volume F(x,t) is included in equation (2.3)

$$\frac{1}{\rho(x)}\nabla p(x,t;x_s) + \frac{\partial u(x,t)}{\partial t} = \frac{1}{\rho(x)}F(x,t); \tag{2.13}$$

b) If mass is introduced into the space at a rate of G(x,t) per unit volume, equation (2.8) becomes

$$\frac{1}{\rho(x)c(x)^2}\frac{\partial p(x,t;x_s)}{\partial t} + \nabla \cdot u(x,t) = \frac{1}{\rho(x)}G(x,t). \tag{2.14}$$

Taking the divergence of equation (2.13), and the partial derivative respect time of equation (2.14), and combining both equations, an inhomogeneous wave equation

is obtained

$$\frac{1}{c(x)^2}\frac{\partial^2 p(x,t;x_s)}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla p(x,t;x_s)\right] = \frac{\partial G(x,t)}{\partial t} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}F(x,t)\right]. \tag{2.15}$$

In the region without external forces, the right-hand side of equation (2.15) vanishes, giving the homogeneous wave equation.

2.2 Theoretical aspects of the standard traveltime inversion

Following the mathematical procedure given by Luo and Schuster (1991), I present in this section the derivation of the wave-equation traveltime inversion method. Let $p(x_r,t;x_s)_{obs}$ and $p(x_r,t;x_s)_{cal}$ denote, respectively, the observed and calculated pressure seismograms at receiver location x_r due to a line source at location x_s. The calculated seismogram $p(x_r,t;x_s)_{cal}$ satisfies the 2D acoustic wave-equation:

$$\frac{1}{c(x)^2}\frac{\partial^2 p(x,t;x_s)}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla p(x,t;x_s)\right] = s(t;x_s), \tag{2.16}$$

where $c(x)$ is the velocity field, $\rho(x)$ is the density function, and $s(t;x_s)$

$$s(t;x_s) = s(t)\delta(x-x_s), \tag{2.17}$$

is the source function.

The one-dimensional cross-correlation function used in the wave-equation traveltime inversion is defined in the time domain by:

$$f(x_r,\tau;x_s) = \int dt \frac{p(x_r,t+\tau;x_s)_{obs}}{A(x_r;x_s)_{obs}} p(x_r,t;x_s)_{cal}, \tag{2.18}$$

with $A(x_r;x_s)_{obs}$ the maximum amplitude of $p(x_r,t;x_s)_{obs}$ and τ the time shift between the observed and the calculated seismograms. This function has a maxi-

mum when the observed and the calculated traces best match each other. It is important to note that each pair of traces is correlated independently, so the moveout is not taken into account in the calculations.

The traveltime residual $\Delta\tau$ that maximises the cross-correlation function is used to define the connective function as

$$\begin{aligned} \dot{f}_{\Delta\tau} &= \left[\frac{\partial f(x_r,\tau;x_s)}{\partial\tau}\right]_{\tau=\Delta\tau} \\ &= \int dt \frac{\dot{p}(x_r,t+\Delta\tau;x_s)_{obs}}{A(x_r;x_s)_{obs}} p(x_r,t;x_s)_{cal} = 0, \end{aligned} \tag{2.19}$$

where $\dot{p} = \partial p(x_r,t;x_s)/\partial t$.

The misfit function S_1, given by

$$S_1 = \frac{1}{2}\sum_s\sum_r \Delta\tau^2(x_r;x_s), \tag{2.20}$$

permits the current velocity model to be corrected by minimizing the delays between the real and the calculated traces. $\Delta\tau(x_r;x_s)$ are the time-shift values where the cross-correlation function, equation (2.18), has a maximum.

The steepest descent method allows the velocity model to be updated

$$c(x)_{k+1} = c(x)_k + \alpha_k \gamma_1(x)_k, \tag{2.21}$$

where $\gamma_1(x)_k$ represents the steepest descent direction of the misfit function S_1, given by

$$\gamma_1(x)_k = -\frac{\partial S_1}{\partial c(x)} = -\sum_s\sum_r \frac{\partial(\Delta\tau)}{\partial c(x)}\Delta\tau(x_r;x_s), \tag{2.22}$$

and α_k is the step length for the kth iteration. α_k is chosen either analytically (Tarantola, 1987) or by trial and error. In practice a preconditioned steepest descent method is used in order to accelerate the convergence.

Using equation (2.19) and the chain rule

$$\frac{\partial \dot{f}}{\partial c(x)} = \frac{\partial \dot{f}}{\partial \Delta\tau}\frac{\partial \Delta\tau}{\partial c(x)}, \tag{2.23}$$

I can write

$$\frac{\partial \Delta\tau}{\partial c(x)} = \frac{\dfrac{\partial \dot{f}}{\partial c(x)}}{\dfrac{\partial \dot{f}}{\partial \Delta\tau}}. \tag{2.24}$$

Hence

$$\frac{\partial(\Delta\tau)}{\partial c(x)} = \frac{1}{E}\int dt \dot{p}(x_r, t+\Delta\tau; x_s)_{obs}\frac{\partial p(x_r, t; x_s)_{cal}}{\partial c(x)}, \tag{2.25}$$

with

$$\begin{aligned} E &= \int dt \ddot{p}(x_r, t+\Delta\tau; x_s)_{obs} p(x_r, t; x_s)_{cal}, \\ &= -\int dt \dot{p}(x_r, t+\Delta\tau; x_s)_{obs} \dot{p}(x_r, t; x_s)_{cal}. \end{aligned} \tag{2.26}$$

In order to calculate $(\partial p(x_r, t; x_s)_{cal}/\partial c(x))$, let's consider the pressure field $p(x, t; x_s)$ that satisfies the wave equation

$$\begin{aligned} \frac{1}{c(x)^2}\frac{\partial^2 p(x,t;x_s)}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla p(x,t;x_s)\right] &= s(t;x_s); \\ p(x,0;x_s) = 0 \qquad \dot{p}(x,0;x_s) &= 0, \end{aligned} \tag{2.27}$$

and the corresponding Green's function $g(x, t; x', t')$

$$\begin{aligned} \frac{1}{c(x)^2}\frac{\partial^2 g(x,t;x',t')}{\partial t^2} - \rho(x)\nabla.\left[\frac{1}{\rho(x)}\nabla g(x,t;x',t')\right] &= \delta(x-x')\delta(t-t'); \\ g(x,t;x',t') = 0, \qquad \dot{g}(x,t;x',t') &= 0, \qquad \text{for } t \le t'. \end{aligned} \tag{2.28}$$

A perturbation $\delta c(x)$ in the velocity field will produce a variation $\delta p(x, t; x_s)$ in

the pressure field which satisfies

$$\frac{1}{[c(x)+\delta c(x)]^2}\frac{\partial^2[p(x,t;x_s)+\delta p(x,t;x_s)]}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla[p(x,t;x_s)+\delta p(x,t;x_s)]\right] = s(t;x_s);$$

$$p(x,0;x_s)+\delta p(x,0;x_s)=0 \qquad \dot{p}(x,0;x_s)+\delta\dot{p}(x,0;x_s)=0, \tag{2.29}$$

Expanding the perturbed velocity function and using equation (2.27), results in

$$\frac{1}{c(x)^2}\frac{\partial^2\delta p(x,t;x_s)}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla\delta p(x,t;x_s)\right] = \frac{\partial^2 p(x,t;x_s)}{\partial t^2}\frac{2\delta c(x)}{c(x)^3}+O(\delta c(x)^2);$$

$$\delta p(x,0;x_s)=0 \qquad \delta\dot{p}(x,0;x_s)=0. \tag{2.30}$$

In terms of the Green's function, the solution of (2.30) is given by (see Morse and Feshbach, 1953)

$$\delta p(x_r,t;x_s) = \int_v dv(x')g(x_r,t;x',0) * \frac{\partial^2 p(x',t;x_s)}{\partial t^2}\frac{2\delta c(x')}{c(x')^3}, \tag{2.31}$$

where $dv(x')$ denotes the volume element at x', and the asterisk indicates time convolution. Equation (2.31) is the so called Fréchet derivative operator.

Integrating by parts the convolutional term in equation (2.31) results in

$$g(x_r,t;x',0) * \frac{\partial^2 p(x',t;x_s)}{\partial t^2} = \int_0^T dt' g(x_r,t-t';x',0)\frac{\partial^2 p(x',t';x_s)}{\partial t^2}$$
$$= g(x_r,t-T;x',0)\dot{p}(x',T;x_s) - g(x_r,t;x',0)\dot{p}(x',0;x_s) - \int_0^T dt'\dot{g}(x_r,t-t',x',0)\dot{p}(x',t';x_s), \tag{2.32}$$

and using the initial conditions given in equations (2.27) and (2.28), I have

$$g(x_r,t;x',0) * \frac{\partial^2 p(x',t;x_s)}{\partial t^2} = -\int_0^T dt' \dot{g}(x_r,t-t';x',0)\dot{p}(x',t';x_s)$$
$$= -\dot{g}(x_r,t;x',0) * \dot{p}(x',t;x_s). \qquad (2.33)$$

Introducing this expression in equation (2.31) yields

$$\delta p(x_r,t;x_s) = -\int_v dv(x')\dot{g}(x_r,t;x',0) * \dot{p}(x',t;x_s)\frac{2\delta c(x')}{c(x')^3}. \qquad (2.34)$$

Considering a point scatterer in the velocity field

$$\delta c(x) = \Delta c\, \delta(x'-x), \qquad (2.35)$$

the perturbation in the pressure field $\delta p(x_r,t;x_s)$ becomes

$$\delta p(x_r,t;x_s) = -\dot{g}(x_r,t;x,0) * \dot{p}(x,t;x_s)\frac{2\Delta c}{c(x)^3}. \qquad (2.36)$$

Dividing by Δc, and using the reciprocity property of the Green's function (Aki and Richards, page 28, 1980), I get

$$\frac{\partial p(x_r,t;x_s)_{cal}}{\partial c(x)} = -\frac{2}{c(x)^3}\dot{g}(x,t;x_r,0) * \dot{p}(x,t;x_s). \qquad (2.37)$$

Returning to equation (2.25) and using the fact that E is not a function of time (see equation(2.26)), I obtain

$$\frac{\partial(\Delta\tau)}{\partial c(x)} = -\frac{1}{c(x)^3}\int dt\dot{g}(x,t;x_r,0) * \dot{p}(x,t;x_s)\frac{2\dot{p}(x_r,t+\Delta\tau;x_s)_{obs}}{E}, \qquad (2.38)$$

and equation (2.22) becomes

$$\gamma_1(x)_k = \frac{1}{c(x)^3}\sum_s\sum_r\int dt\dot{g}(x,t;x_r,0) * \dot{p}(x,t;x_s)\delta\tau(x_r,t;x_s), \qquad (2.39)$$

where $\delta\tau(x_r, t; x_s)$ is defined as the pseudo-traveltime residual .

$$\delta\tau(x_r, t; x_s) = \frac{2}{E}\dot{p}(x_r, t + \Delta\tau; x_s)_{obs}\Delta\tau(x_r; x_s). \tag{2.40}$$

Rewriting equation (2.39) as

$$\gamma_1(x)_k = \frac{1}{c(x)^3}\sum_s\sum_r \int dt\delta\tau(x_r, t; x_s)\int dt'\dot{g}(x, t - t'; x_r, 0)\dot{p}(x, t'; x_s), \tag{2.41}$$

and exchanging the integrals, yields

$$\gamma_1(x)_k = \frac{1}{c(x)^3}\sum_s\sum_r \int dt'\dot{p}(x, t'; x_s)\int dt\dot{g}(x, t - t'; x_r, 0)\delta\tau(x_r, t; x_s). \tag{2.42}$$

Now, using the fact that the Green's function is invariant to time translation (Aki and Richards, 1980), yields

$$\begin{aligned}
\gamma_1(x)_k &= \frac{1}{c(x)^3}\sum_s\sum_r \int dt'\dot{p}(x, t'; x_s)\int dt\dot{g}(x, 0; x_r, t' - t)\delta\tau(x_r, t; x_s), \\
&= \frac{1}{c(x)^3}\sum_s\sum_r \int dt'\dot{p}(x, t'; x_s)\dot{g}(x, 0; x_r, t') * \delta\tau(x_r, t'; x_s), \\
&= \frac{1}{c(x)^3}\sum_s\sum_r \int dt'\dot{p}(x, t'; x_s)\dot{g}(x, -t'; x_r, 0) * \delta\tau(x_r, t'; x_s).
\end{aligned} \tag{2.43}$$

Defining $q_1(x, t; x_s)$ as

$$q_1(x, t; x_s) = \sum_r g(x, -t; x_r, 0) * \delta\tau(x_r, t; x_s), \tag{2.44}$$

and rewriting it as

$$q_1(x, t; x_s) = \sum_r \int dt' g(x, 0; x_r, t - t')\delta\tau(x_r, t'; x_s), \tag{2.45}$$

I have,

$$\begin{aligned}\dot{q}_1(x,t;x_s) &= \frac{\partial}{\partial t}\sum_r \int dt' g(x,t'-t;x_r,0)\delta\tau(x_r,t';x_s) \\ &= -\sum_r \int dt' \dot{g}(x,t'-t;x_r,0)\delta\tau(x_r,t';x_s) \\ &= -\sum_r \int dt' \dot{g}(x,0;x_r,t-t')\delta\tau(x_r,t';x_s) \\ &= -\sum_r \dot{g}(x,0;x_r,t) * \delta\tau(x_r,t;x_s). \end{aligned} \tag{2.46}$$

Combining equation (2.46) with equation (2.43), results in

$$\gamma_1(x)_k = -\frac{1}{c(x)_k^3}\sum_s \left[\int dt\, \dot{p}(x,t;x_s)\dot{q}_1(x,t;x_s)\right]. \tag{2.47}$$

In the previous equations, $q_1(x,t;x_s)$ represents the missing pressure field, and is computed by reverse time propagation of the pseudo-traveltime residual $\delta\tau(x_r,t;x_s)$ acting as a source at the receiver location x_r.

The wave-equation traveltime inversion reduces to full waveform inversion when the current model is close to the real model. Also, it is reduced to ray-tracing traveltime tomography in the high frequency approximation (Luo and Schuster, 1991).

2.3 Numerical examples and discussion

A computer code for full waveform inversion was provided by the University of Utah Tomography Consortium, and I modified it in order to execute the traveltime inversion. I use the resulting standard wave-equation traveltime inversion code to invert different datasets. The first one corresponds to the synthetic crosshole dataset used by Luo and Schuster (1991), and by Zhou et al. (1995) in their papers. I use this synthetic dataset mainly to verify that the modified code works properly and to identify the differences between the inversion of crosshole datasets and surface reflection datasets. The main example for the surface reflection geometry is a point-diffractor embedded in a constant background velocity of 2000 m/s.

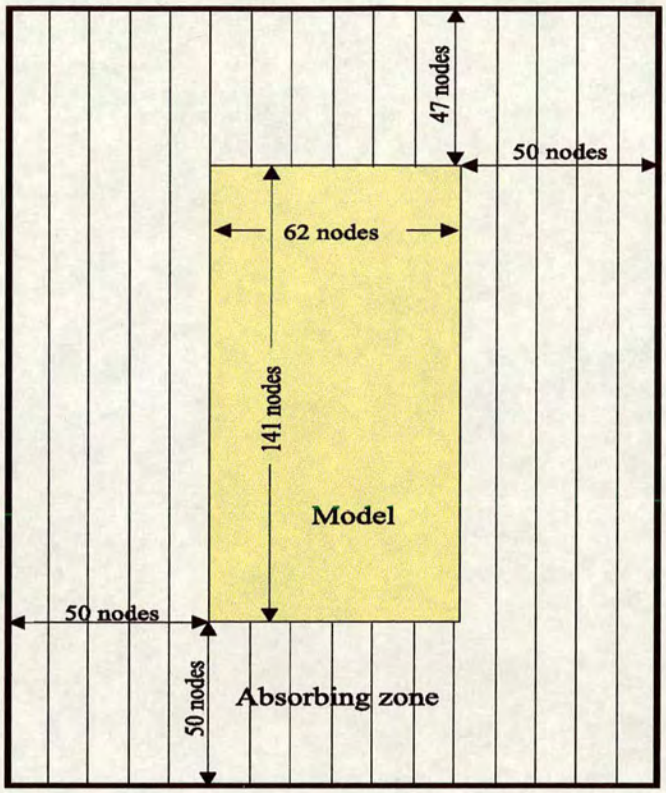
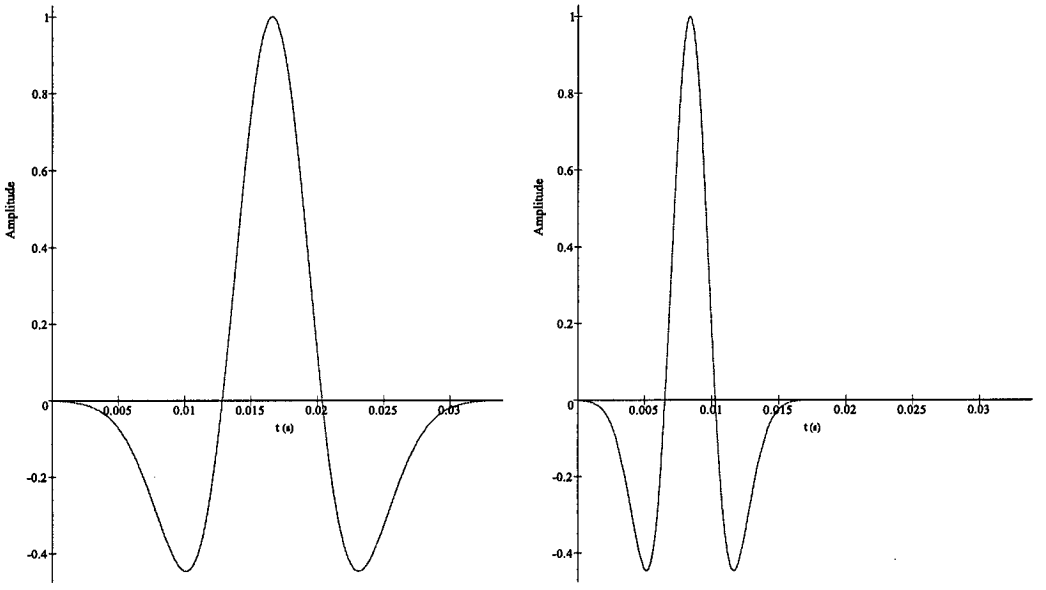


Figure 2.3: Grid.

The grid (Figure 2.3) consists of 62x141 nodes with an evenly spaced distance of 1.5 m. Four absorbing boundaries are considered. A Ricker wavelet (Sheriff and Geldart, 1995) having a peak frequency of 60 Hz is used for modelling and inverting the crosshole dataset, while 120 Hz is used for the surface reflection datasets, see Figure 2.4. The sample rate in time is 0.2 msec. and the step length is kept constant during the inversion. The density function (Gardner et al., 1974; Luo and Schuster, 1991) is given by

$$\rho(x) = \rho_0[c(x)/c_{p00}]^{den}, \tag{2.48}$$

with the values ρ_0=2.1 Kg/m^3, c_{p00}=6500 m/s and den=1/2 provided by the Utah Tomography Consortium for the synthetic crosshole dataset. Regardless of the geometry configuration, the amplitudes of the seismograms are normalized



a) Peak frequency: 60 Hz b) Peak frequency: 120 Hz

Figure 2.4: Ricker wavelets.

prior to muting out any waveform, i.e., each trace is normalized to its maximum amplitude value.

Even though the delays in time can be calculated using the cross-correlation function, the program uses a simpler procedure for the crosshole configuration: the traveltimes t_{cal} and t_{obs} corresponding to the maximum amplitude for each trace in the calculated and the observed seismograms, respectively, are determined. Then, the time shift $\Delta\tau$ is given simply by the subtraction of these two values. In this particular case these values are associated with the transmitted arrivals.

For surface reflection datasets, the one-dimensional cross-correlation function is computed using the Fast Fourier transform (Bracewell, 1986; Press et al., 1994; Brook and Wynne, 1988). The observed and the calculated seismograms are Fourier transformed to the frequency domain, multiplied in that domain, and inverse Fourier transformed back to the time domain. This process is represented

by:

$$p(x_r, t; x_s)_{obs} \otimes p(x_r, t; x_s)_{cal} \Longleftrightarrow P(x_r, \omega; x_s)_{obs} \tilde{P}(x_r, \omega; x_s)_{cal}, \qquad (2.49)$$

where $\otimes$ means cross-correlation, tilde denotes complex conjugate, and P_{obs} and P_{cal} are the Fourier transforms of p_{obs} and p_{cal}, respectively. The time shifts are the delays associated with the maximum amplitudes in the cross-correlogram.

The geometry for the crosshole configuration consists of 15 shots regularly spaced along the left-hand side of the model and 36 receivers distributed along the right-hand side. The complete seismograms are used for the inversion. Figure 2.5a shows the original synthetic model and Figure 2.5b shows the result of the standard traveltime inversion after 10 iterations, starting with a 3000 m/s constant velocity model. The resulting traveltime tomogram contains the long wavelength features of the original velocity model.

The next step is to apply the full waveform inversion (explained in Chapter 3) to recover the high frequencies present in the original model. The output from the traveltime inversion is used as the starting velocity field. The strong direct arrivals resolve the low wavenumber velocity variations of the part of the model they pass through. In order to account for the high wavenumber details of the model, the reflection events have to be considered. So, the direct arrivals should be muted out in the observed and the calculated seismograms.

Four additional iterations are performed using the full waveform inversion. The result displayed in Figure 2.5c is comparable to the result given by Zhou et al. (1995), and the differences can be related to the fact they used different parameters for the inversion, e.g., the grid size, the peak frequency of the source, the number of shots, etc. Also, they removed all waveforms except the transmitted arrivals when using the traveltime inversion. This result suggests the computer code is working properly.

The surface reflection geometry is simulated by spreading 31 receivers and 15

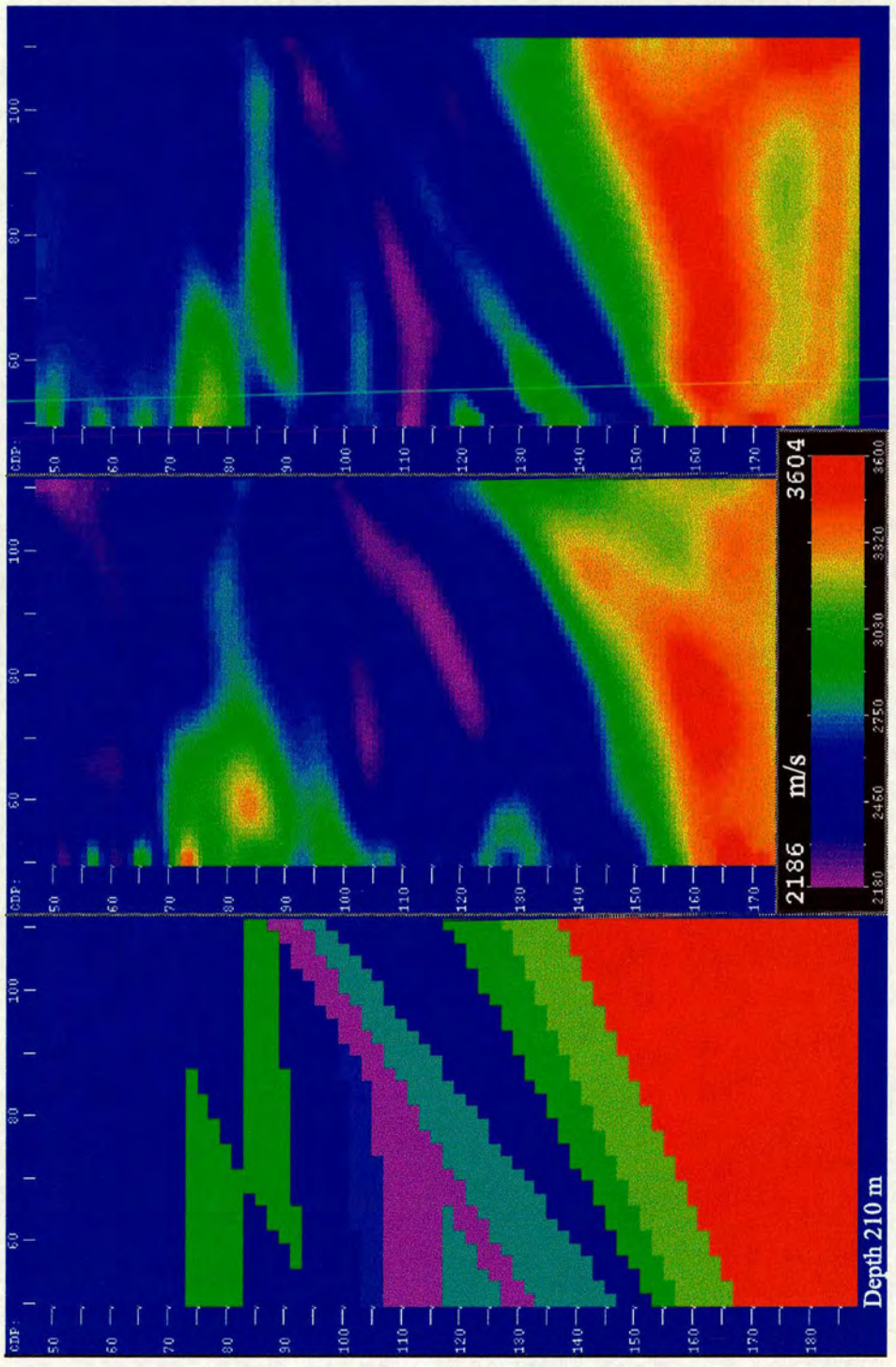
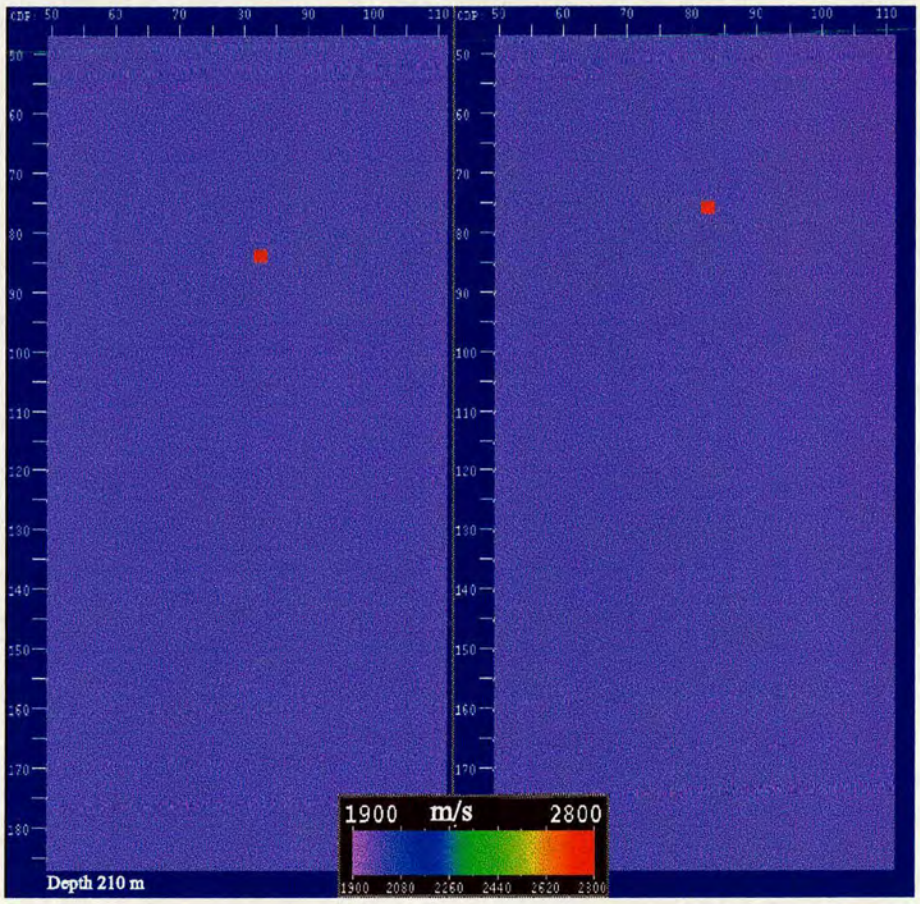


Figure 2.5: Crosshole dataset inversion: a) Synthetic model, b) Standard traveltime inversion, c) Full waveform inversion after traveltime inversion.

shots along the surface. There is a shot every 4th node and one receiver every other node. The initial velocity model (Figure 2.6) is a point-diffractor located in a position shallower than the original one and with the same background velocity. In surface reflection datasets the first arrivals provide only near-surface information, then the strong direct arrivals should be muted out prior to cross-correlation and reverse-time propagation. Figure 2.7 displays the results of the traveltime inversion after 5, 10, and 15 iterations and Figure 2.8 depicts the corresponding seismograms for the shot located at the left-hand side of the model.



a) Original velocity model **b) Starting velocity model**

Figure 2.6: Velocity models for traveltime inversion considering surface reflection geometry.

It is clear in this example that the depth of the point-diffractor is compensated

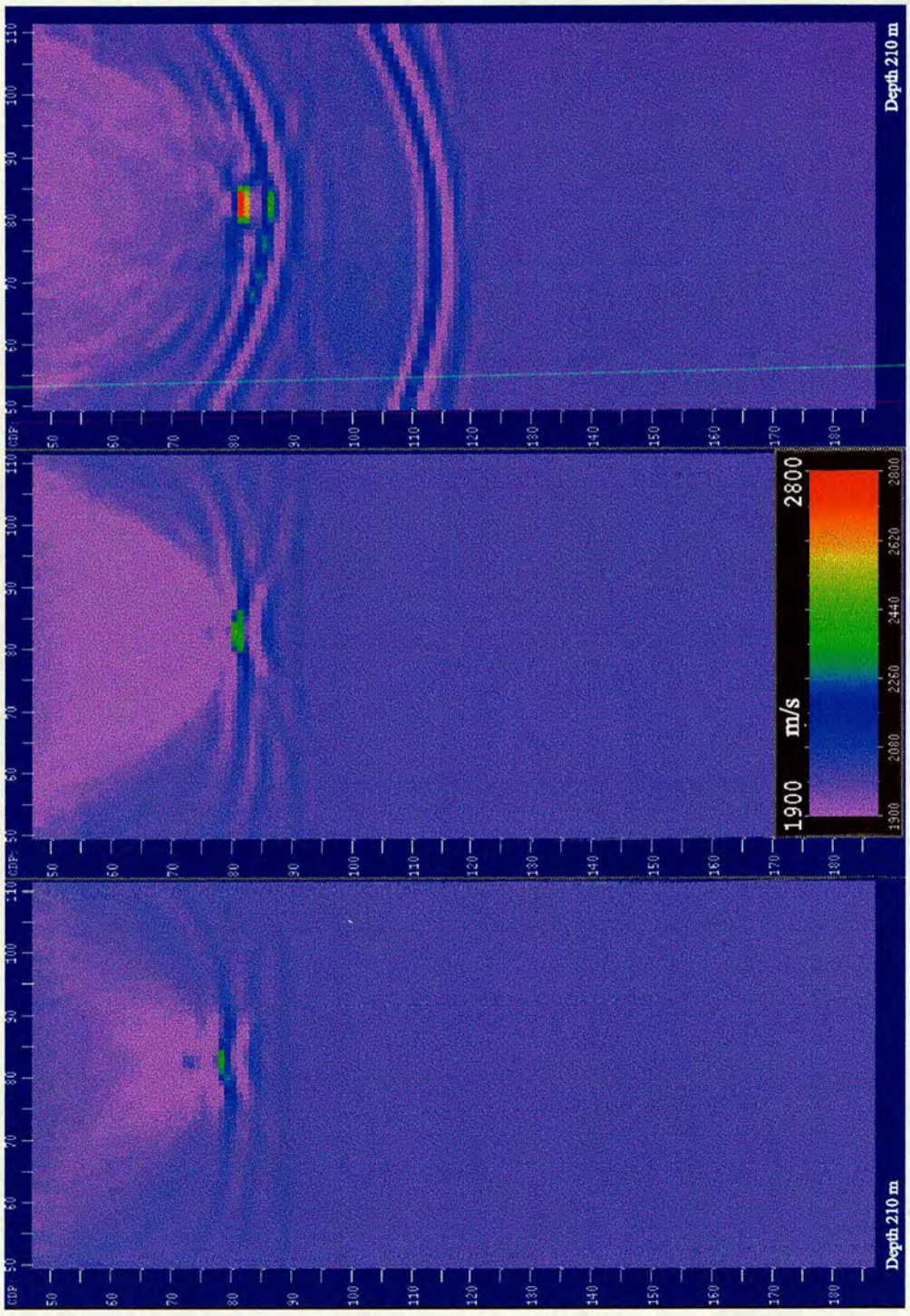


Figure 2.7: Standard traveltime inversion of the point-diffractor model.

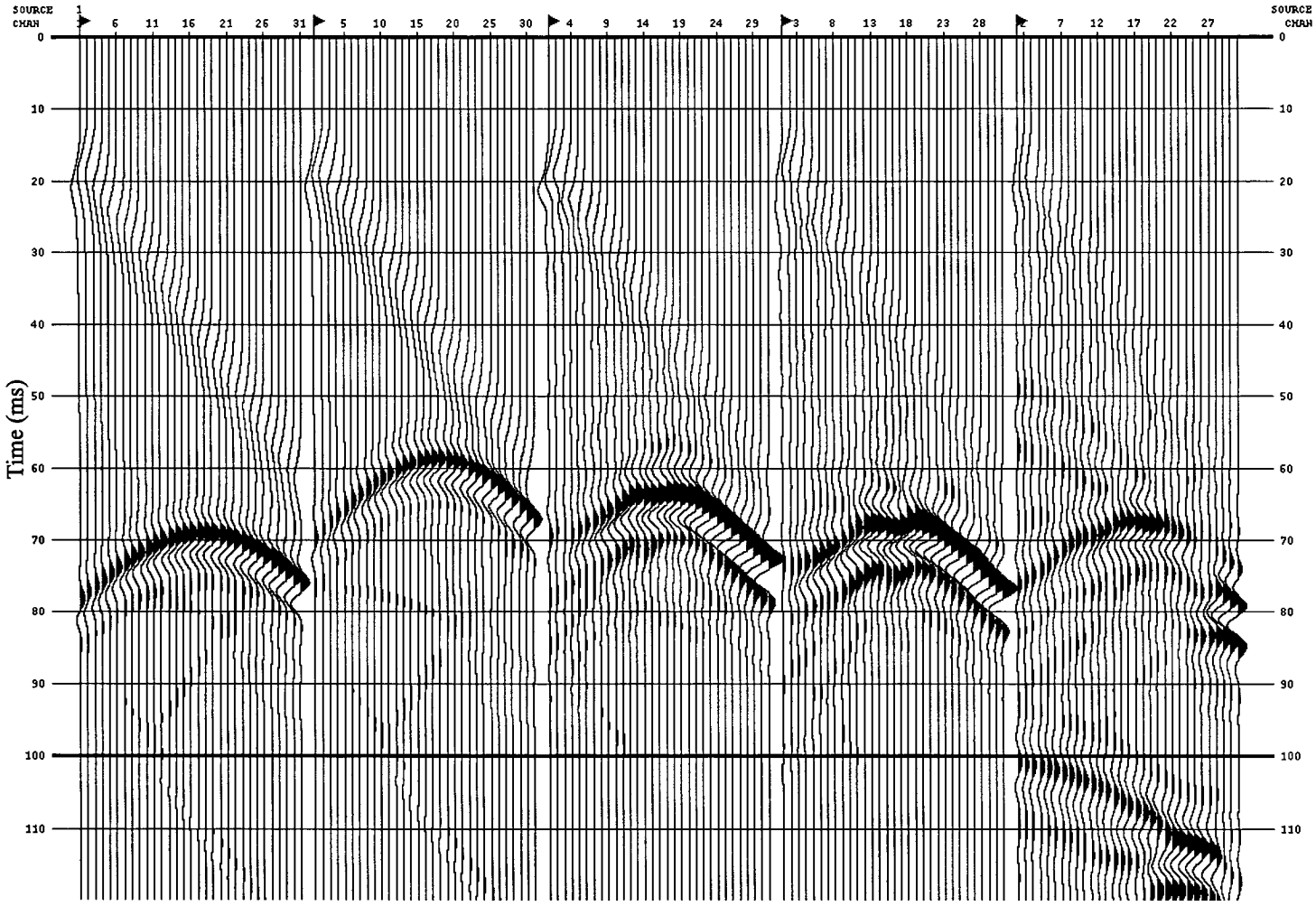


Figure 2.8: Seismograms from the point-diffractor traveltime inversion.

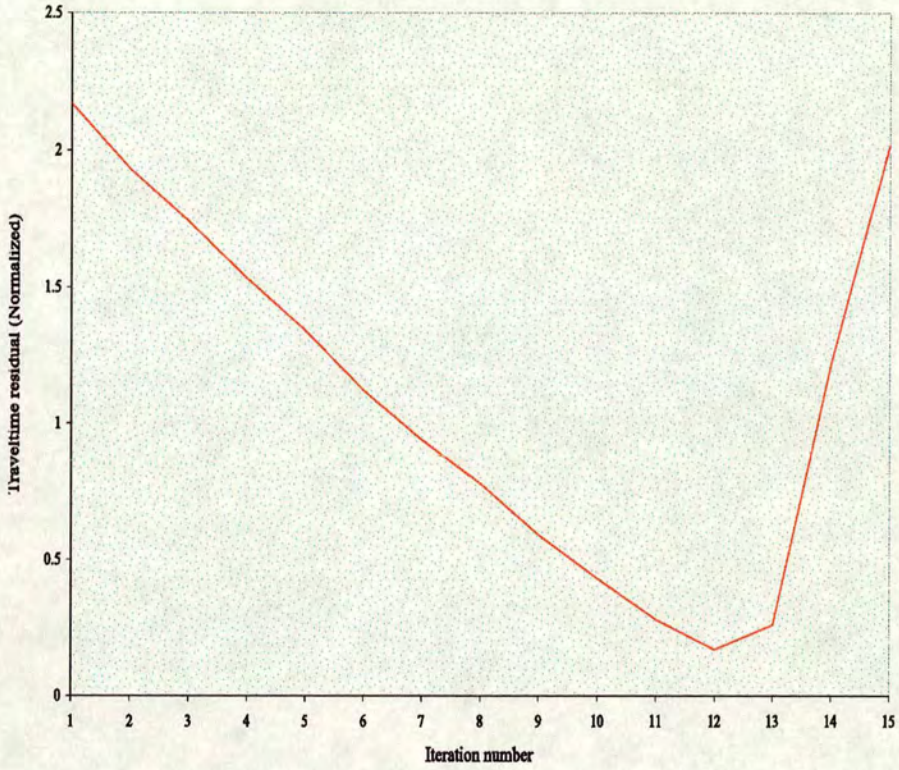


Figure 2.9: Residuals of the point-diffractor traveltime inversion.

by diminishing the velocity above it, until the minimum velocity allowed in the inversion is reached. The minimum velocity is constrained to be 1900 m/s. Although the match between the calculated and the observed seismograms improves in successive iterations, some artifacts appear after the 13th iteration. Then, the traveltime residual jumps to a higher value, as shown in Figure 2.9, and the inversion diverges.

The behaviour of the time delays for each shot gather as a function of offset is shown in Figures 2.10, 2.11, and 2.12 at iterations 1, 12, and 13. In these figures every colour line with the same symbol represents a common shot gather. As each trace is shifted in time independently, the moveout in the observed shot gather is not preserved. The inversion diminishes the traveltime delays until the

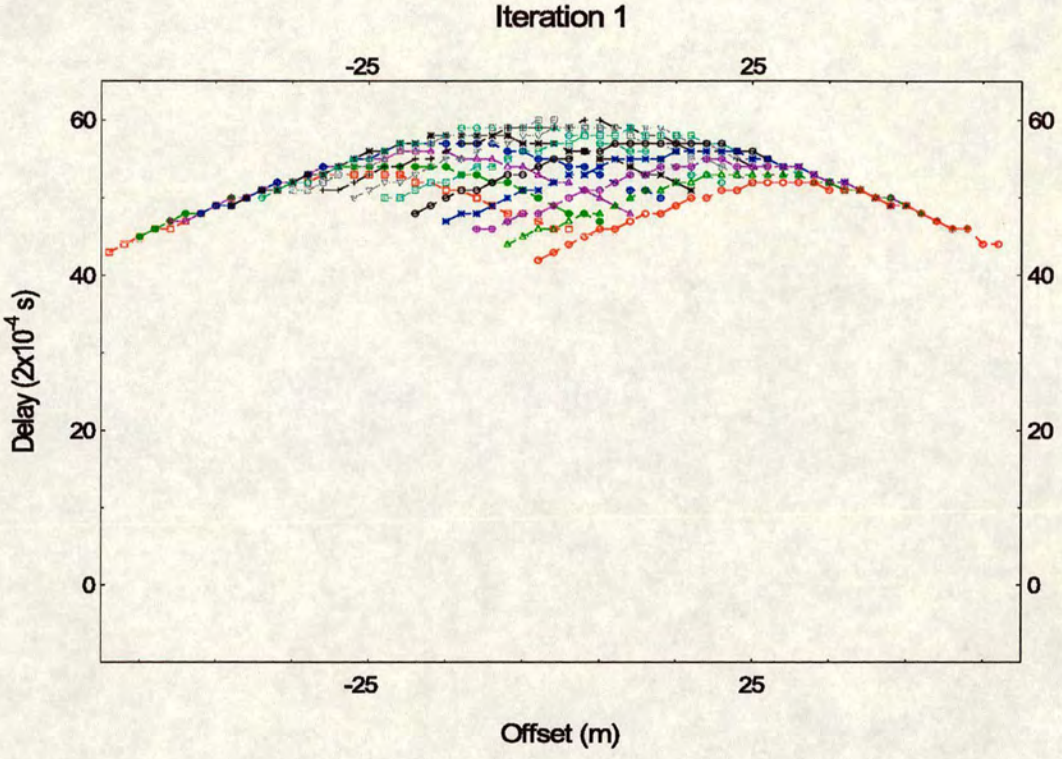


Figure 2.10: Delays associated with iteration 1.

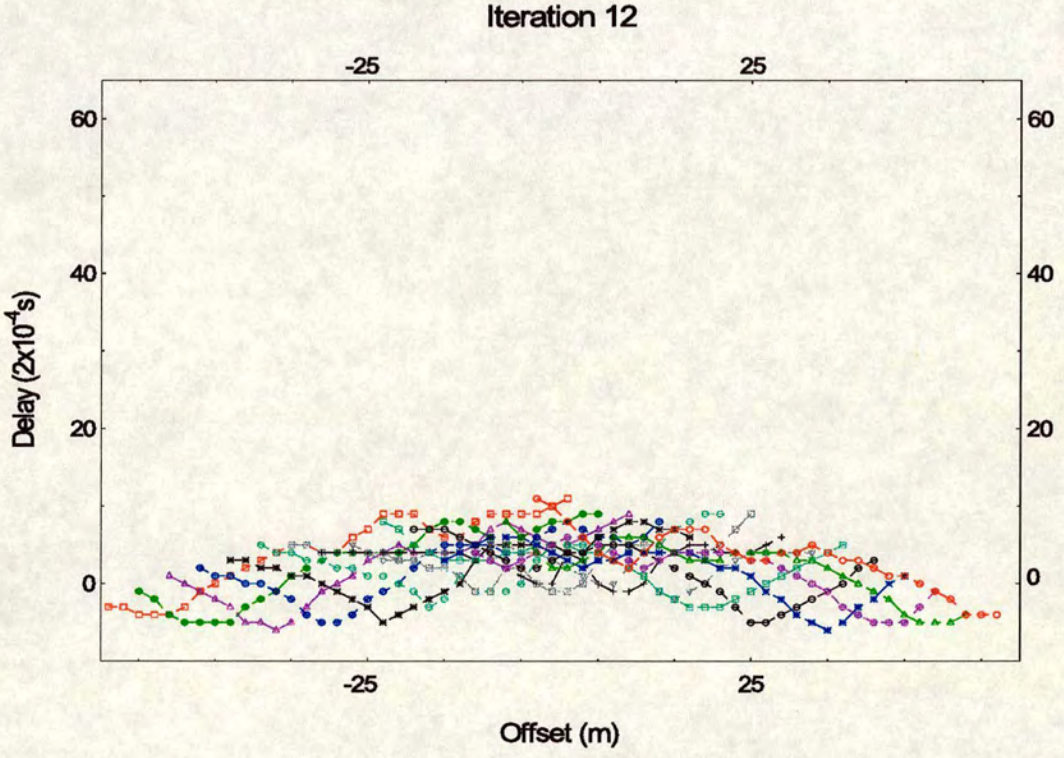


Figure 2.11: Delays associated with iteration 12.

12th iteration. At iteration 13 the inversion starts to diverge and the artifacts predominate over the diffraction. In particular, Figure 2.13 depicts the delays for shot 8 as the iterations proceed. The delays decrease regularly until iteration number 12, and after that the amplitude of the artifacts in the calculated shot gathers governs the traveltime delay.

The artifacts can be controlled by adjusting the step length. The result using a variable step length α_k is shown in Figure 2.14 after 23 iterations. α_k is calculated by making it half of the previous value when the traveltime residual at iteration $k+1$ is bigger than the residual at iteration k (Tarantola, 1984). The traveltime residuals and the step length values can be seen in Figure 2.15. The diffractor is placed one node above the right position due to the lower velocity in the upper part of the model. Some artifacts are still present, and a stable convergence is not reached.

Many parameters are involved in the procedure and all of them play an important role. Here, I present two of them: the minimum velocity allowed during the inversion and, secondly, the size of the model. In Chapter 3, I show the full waveform inversion using two different wavelets as well as the inversion using a model with two different vertical lengths. In Chapter 6 other parameters are considered. Figure 2.16 displays the result for the same point-diffractor model but using 2000 m/s as the minimum velocity allowed in the inversion. After 35 iterations, the diffractor is located at the right depth but some artifacts are still present. The time residual at iteration 35 is not null (see Figure 2.17), and a slight change in the step length value can cause the divergence of the inversion.

Two models are now considered which consist of 2 point-diffractors in a medium with 2000 m/s as background velocity. Again the minimum velocity is 2000 m/s but the total depth is 210 m for the first model and 91.5 m for the second one. The grid spacing remains unaltered at 1.5 m, and only the number of nodes in depth is reduced. Figure 2.18 shows that the artifacts are present in both models but they are extended all-around the model in the second one. Even though the original

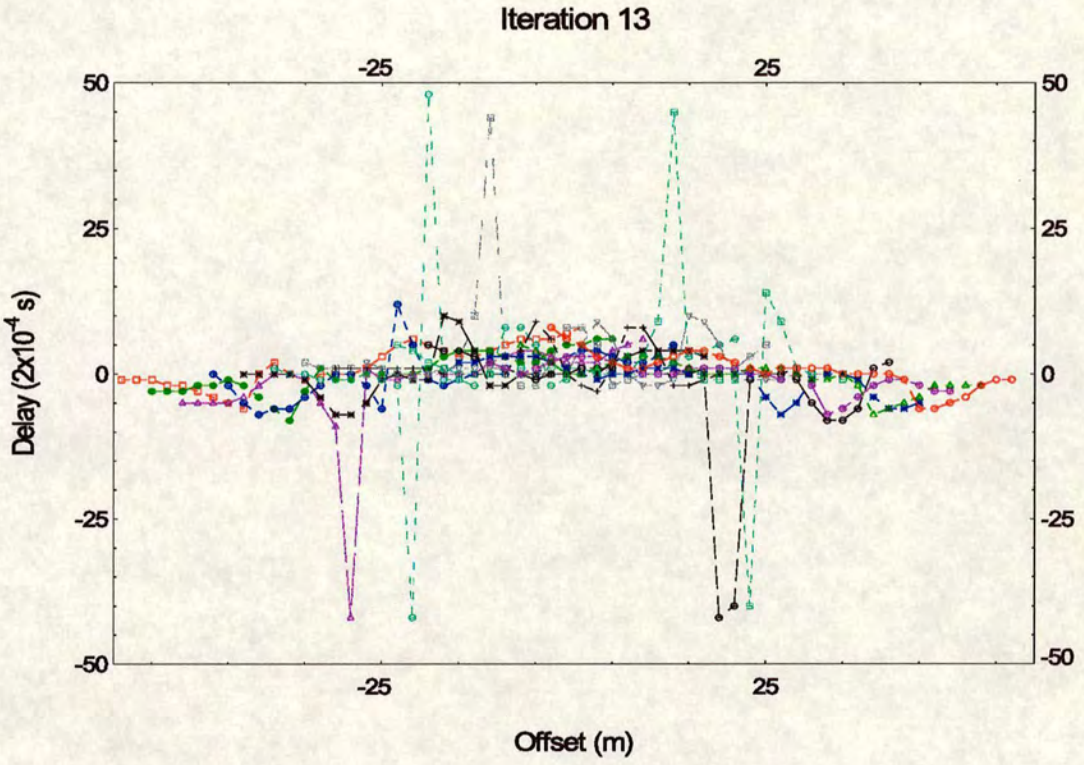


Figure 2.12: Delays associated with iteration 13.

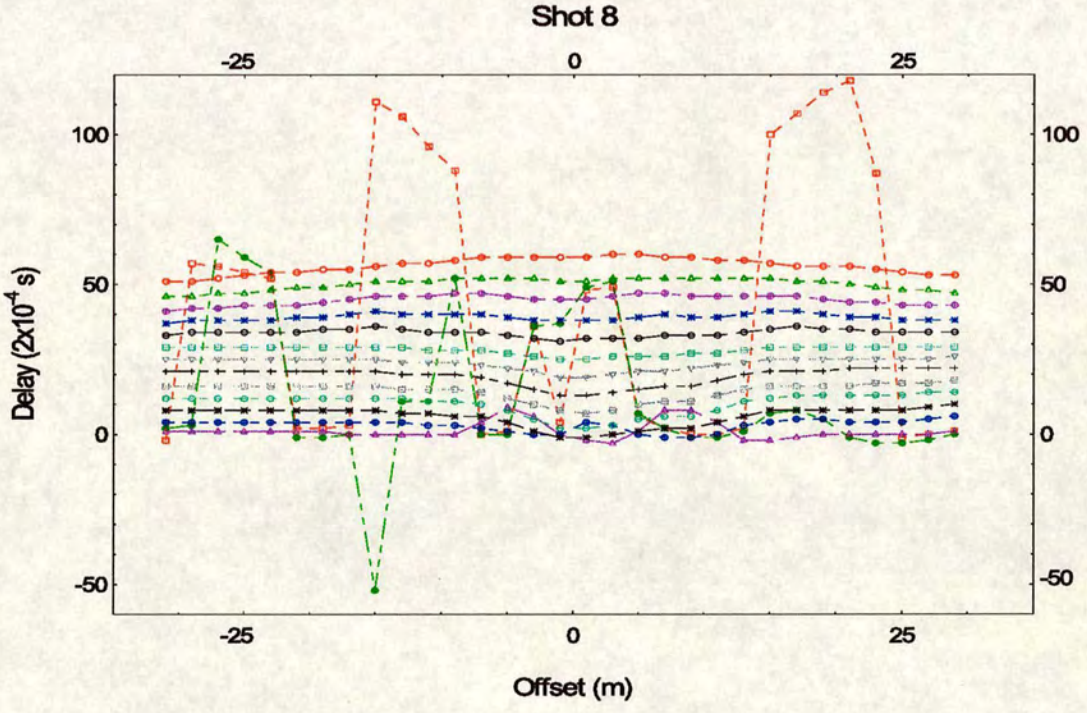
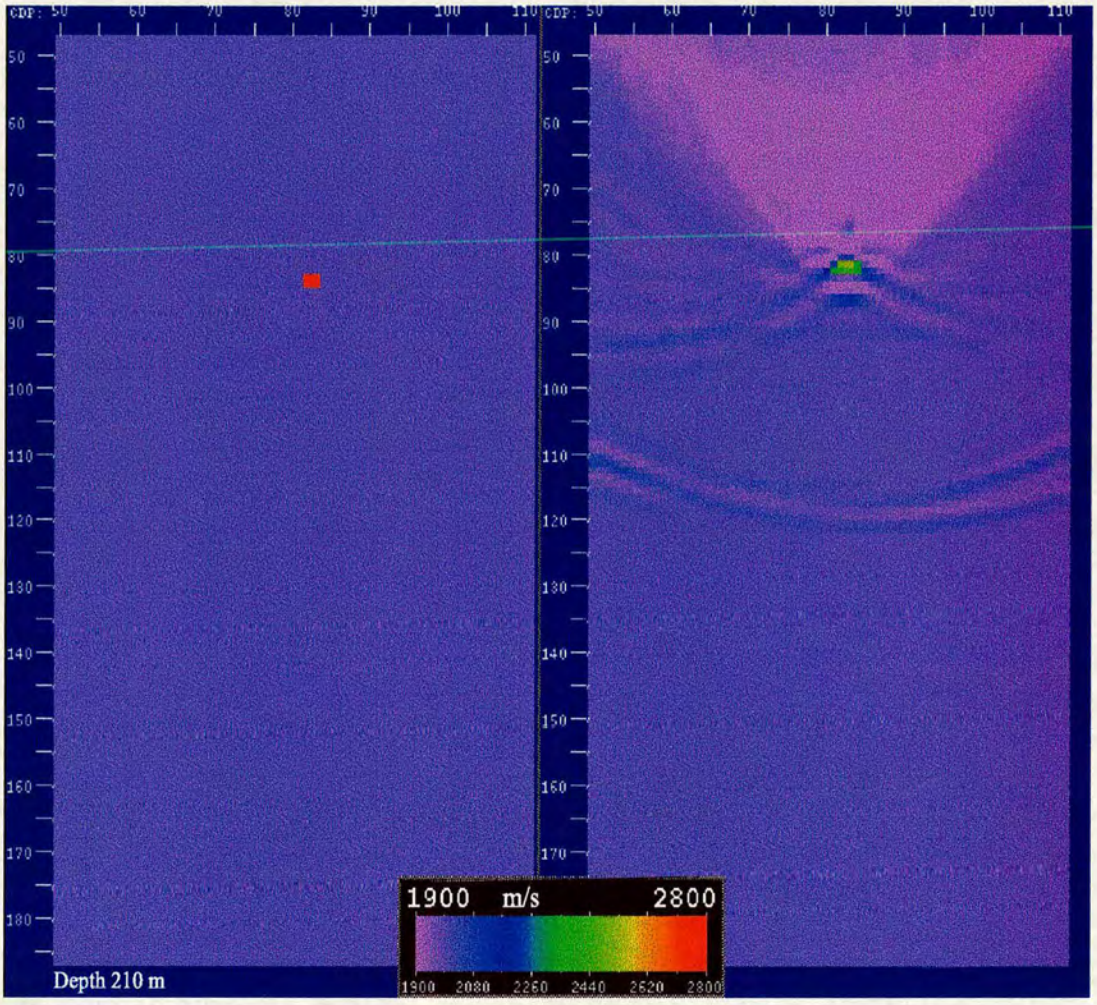


Figure 2.13: Delays associated with shot 8 in 15 iterations.



a) Original model **b) After 23 iterations**

Figure 2.14: Inversion of the point-diffractor using variable step length.

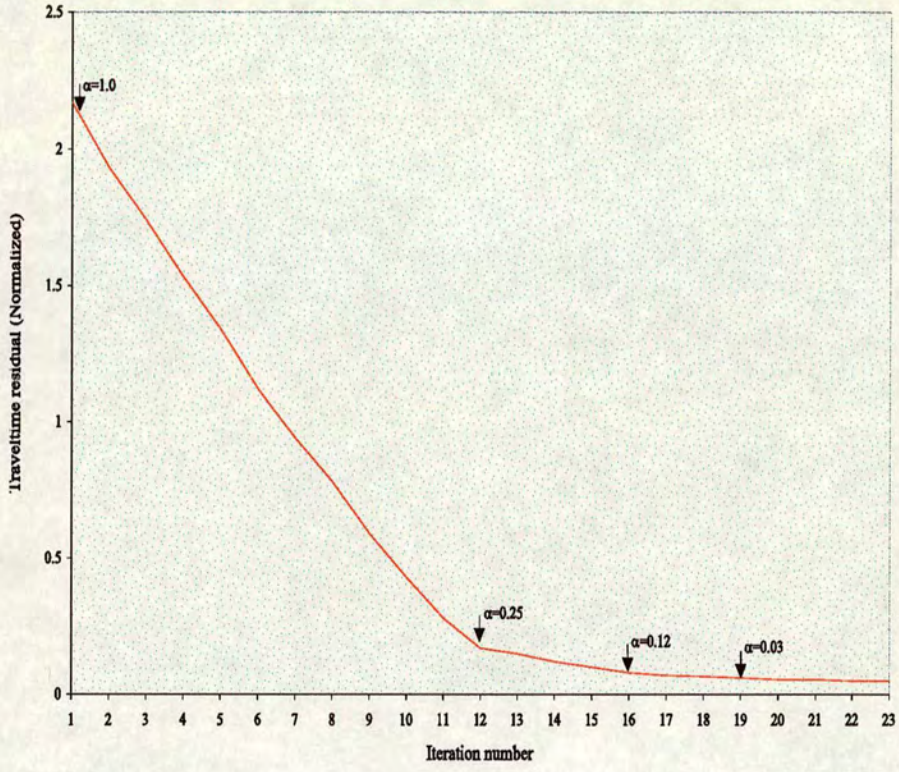


Figure 2.15: Residuals using variable step length.

diffractors and the initial diffractor have the same position in both models, the behaviour is more erratic in the shallower model. This fact may suggest that weak reflections in the lower absorbing boundary are affecting the results, but I show in Chapter 6 that this effect is not seen in the traveltime-offset inversion even though the same wavefield modelling program is used.

2.4 Summary

The wave-equation traveltime inversion provides the low frequency content of the model when the transmitted arrivals are used in the crosshole configurations. Although there are no restrictions in the theory involved in the procedure, the method fails to reconstruct models with surface reflection geometries. In practice,

there are differences in the procedure for each configuration. Among these differences the most relevant are: a) the propagation of the wave field in the model and the predominance of different waveforms due to the arrangement of the sources and receivers, b) the mute of different waveforms present in the shot gathers and c) the amplitude associated with each event. However, in theory, none of them represent a constraint that limits the convergence of the traveltime inversion.

When processing surface reflection datasets, the step length can help to control the influence of the artifacts. However, the convergence of the standard traveltime inversion is unstable. Different results are obtained when the lower boundary of the model is placed at different depths, and the velocity-depth ambiguity is not solved by this inversion method.

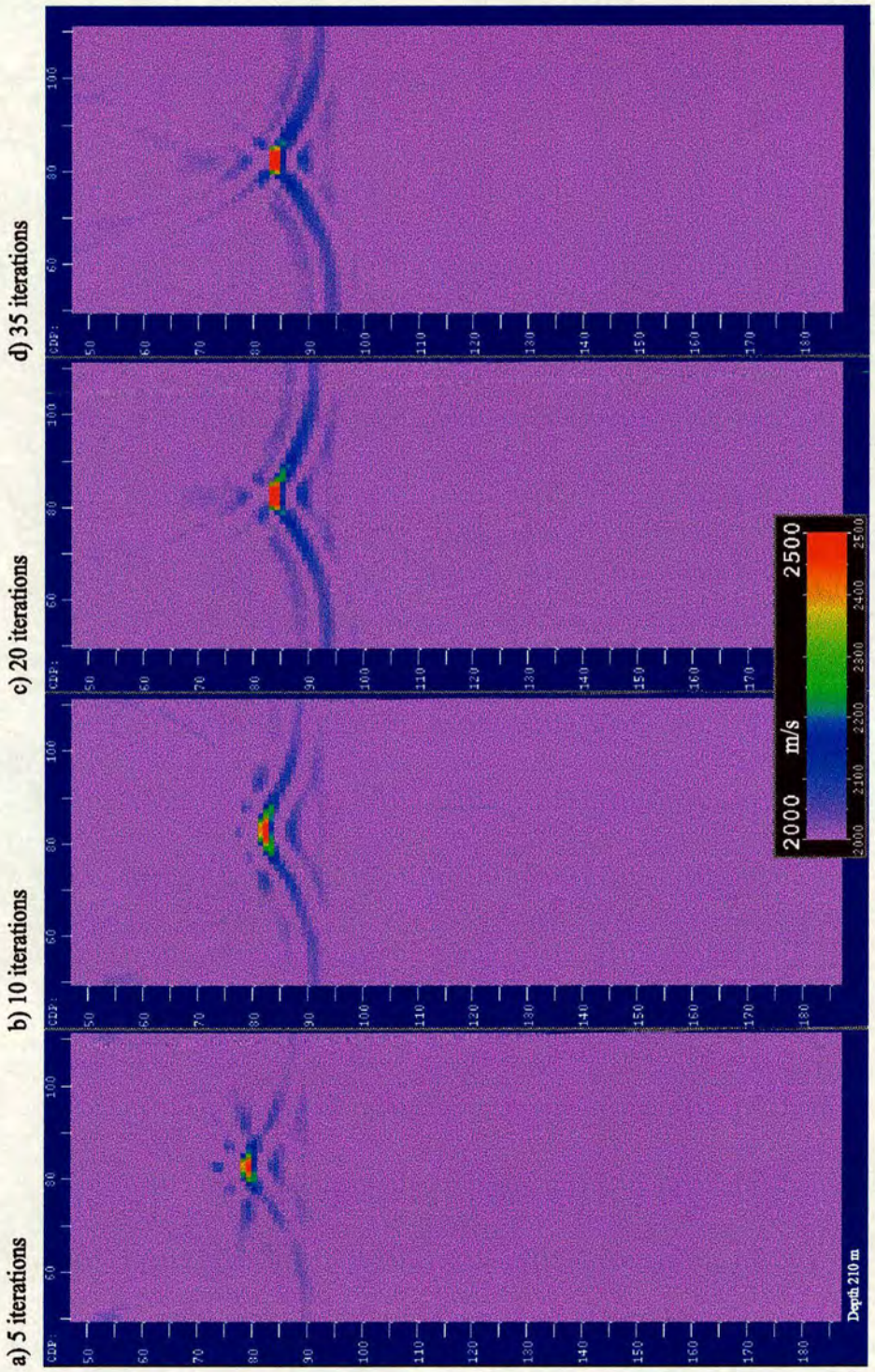


Figure 2.16: Standard traveltime inversion with the minimum velocity constrained to 2000 m/s.

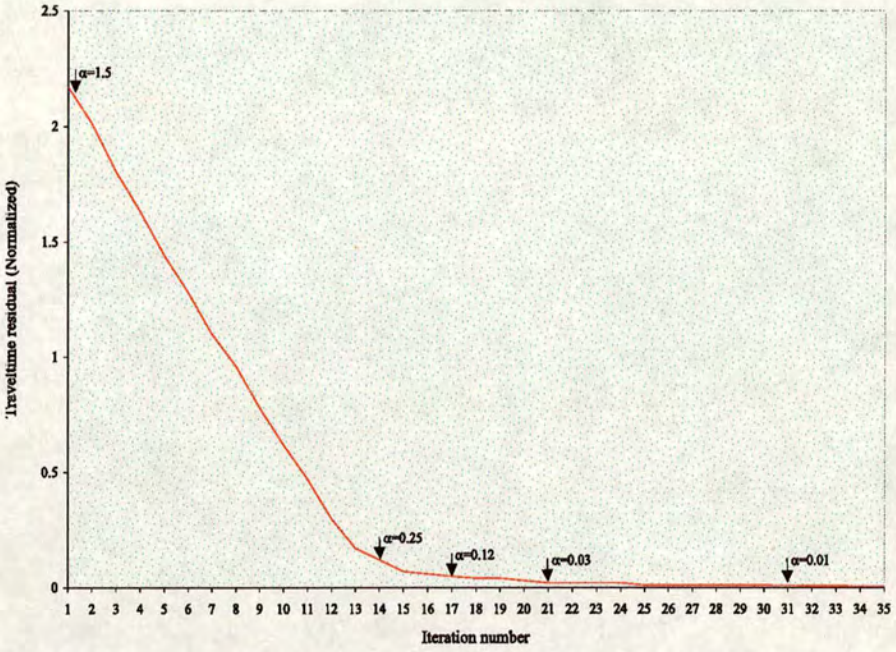


Figure 2.17: Residuals using 2000 m/s as the minimum velocity.

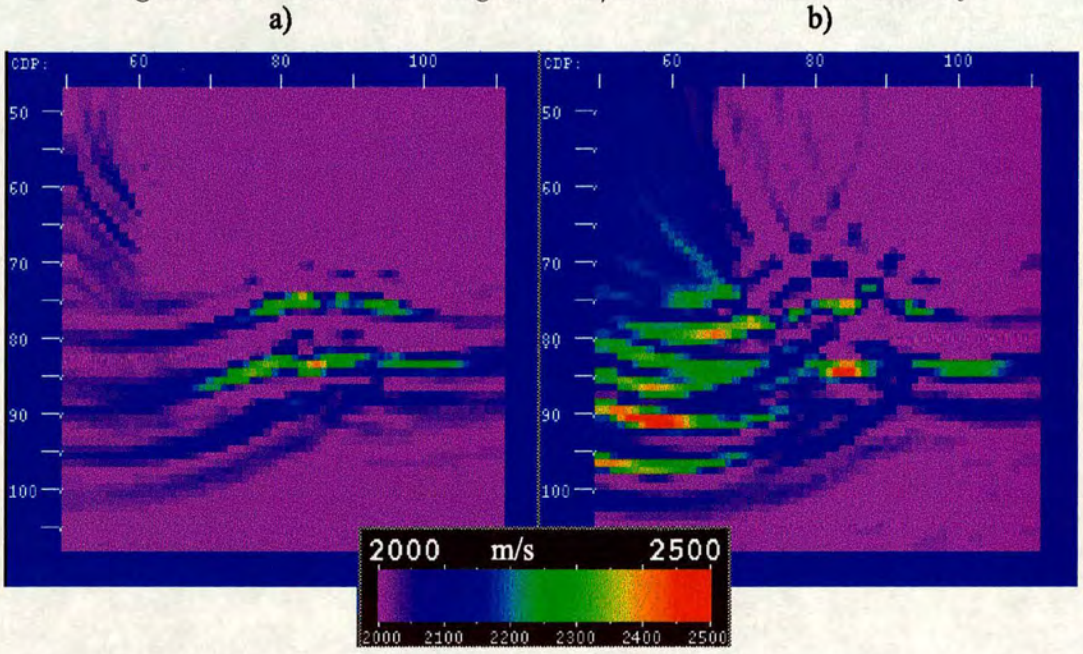


Figure 2.18: Comparison using the same model with two different depths. a) Maximum depth: 210 m, b) Maximum depth: 90.5 m.

Chapter 3

Full waveform inversion

The full waveform inversion has been used to reconstruct the high frequency features present in the original velocity model. This method has the advantage that no high frequency approximations are necessary, but it requires an initial velocity model containing the long wavelength structures of the real model in order to avoid getting stuck in a local minimum. Therefore, an output from the traveltime inversion, or from the traveltime-offset inversion, is a good choice as a starting model for the full waveform inversion. This idea was proposed by Zhou et al. in 1995. They combined the wave-equation traveltime inversion with the full waveform inversion in a hybrid method that allowed them to keep the advantages of both methods. Following this idea, I introduce in Chapter 4 the traveltime-offset inversion as a hybrid method instead of using the one-dimensional traveltime inversion.

I apply the full waveform inversion after the standard traveltime inversion in Chapter 2, and I combine it with the new traveltime-offset inversion method in Chapter 6. Here, in Chapter 3, I present a description of the theory involved in this method together with synthetic examples, and I show how some parameters can affect the results of this inversion method.

3.1 Theory

The general approach for full waveform inversion was proposed by Tarantola in 1984. Gauthier et al. (1986), proved the feasibility of the method by applying it to synthetic datasets. Since then, several authors have used the strengths of this method to reconstruct the fine details present in the velocity models. In this

section, I present a review of the theoretical ideas of the full waveform inversion. I follow the conventions used by Tarantola (1987) with the difference that I use the velocity field $c(x)$ instead of the bulk modulus model K.

Let $\mathbf{p}$ and $\mathbf{C}_p$ represent the dataset and the covariance operator respectively. Here, $\mathbf{C}_p$ describes the experimental uncertainties, and let us assume it is diagonal

$$C_p(x_r, t; x_s; x'_r, t'; x'_s) = \sigma^2_{sd}(x_r, t; x_s)\delta^{rr'}\delta(t - t')\delta^{ss'}, \tag{3.1}$$

where σ^2_{sd} is the standard deviation, and δ^{ab} represents the Krönecker's symbol (1 if $a = b$; 0 otherwise). The pressure wavefield associated with the velocity model $c(x)$ is denoted by

$$\mathbf{p} = \mathbf{f}(\mathbf{c}). \tag{3.2}$$

The observed pressure field is represented as $\mathbf{p}_{obs}$, so

$$\delta\mathbf{p} = \mathbf{f}(\mathbf{c}) - \mathbf{p}_{obs}. \tag{3.3}$$

Defining the norm $\|\delta\mathbf{p}\|$ as

$$\begin{aligned} \|\delta\mathbf{p}\|^2 &= <\delta\hat{\mathbf{p}}, \delta\mathbf{p}> = <\mathbf{C}_p^{-1}\delta\mathbf{p}, \delta\mathbf{p}> = \delta\mathbf{p}^t\mathbf{C}_p^{-1}\delta\mathbf{p} \\ &= \sum_s \int dt \sum_r \delta\hat{p}(x_r, t; x_s)\delta p(x_r, t; x_s), \end{aligned} \tag{3.4}$$

the misfit function S_2, given by the square norm of $\delta\mathbf{p}$, can be written as

$$\begin{aligned} S_2 &= \frac{1}{2}\|\delta\mathbf{p}\|^2 = \frac{1}{2}\|\mathbf{f}(\mathbf{c}) - \mathbf{p}_{obs}\|^2 \\ &= \frac{1}{2}\left[\mathbf{f}(\mathbf{c}) - \mathbf{p}_{obs}\right]^t \mathbf{C}_p^{-1}\left[\mathbf{f}(\mathbf{c}) - \mathbf{p}_{obs}\right]. \end{aligned} \tag{3.5}$$

The best model is achieved when the misfit function S_2 has a minimum. If $\mathbf{c}_n$

is the best model, the gradient of S_2 will vanish

$$\left[\frac{\partial S_2}{\partial \mathbf{c}}\right]_{\mathbf{c}_n} = 0. \tag{3.6}$$

Defining $\mathbf{G}_c$ as

$$\mathbf{G}_c = \frac{\partial \mathbf{f}(\mathbf{c})}{\partial \mathbf{c}}, \tag{3.7}$$

results from equation (3.5)

$$\hat{\gamma}_2 = \frac{\partial S_2}{\partial \mathbf{c}} = \mathbf{G}_c^t \mathbf{C}_p^{-1}[\mathbf{f}(\mathbf{c}) - \mathbf{p}_{obs}]. \tag{3.8}$$

The problem of finding the velocity model that minimises S_2 can be solved using a gradient method. The simplest one is the steepest descent algorithm. Using this method, the new velocity model at iteration $k+1$ is given by

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \alpha_k \gamma_2, \tag{3.9}$$

where α_k is the step length and γ_2 is the steepest descent direction of the misfit function S_2. The direction of the steepest descent is

$$\gamma_2 = -\mathbf{S}_0 \hat{\gamma}_2, \tag{3.10}$$

with $\mathbf{S}_0 = \mathbf{I}$ the preconditioning operator for the steepest descent method, and $\hat{\gamma}_2$ the gradient of the misfit function defined in equation (3.8). Using equation (3.8) to rewrite equation (3.9), I have

$$\begin{aligned} \mathbf{c}_{k+1} &= \mathbf{c}_k - \alpha_k \mathbf{G}_c^t \mathbf{C}_p^{-1}[\mathbf{f}(\mathbf{c}) - \mathbf{p}_{obs}] \\ &= \mathbf{c}_k - \alpha_k \mathbf{G}_c^t \delta\hat{\mathbf{p}}. \end{aligned} \tag{3.11}$$

The Fréchet derivative operator is a linear operator $\mathbf{G}_c$ that associates any

velocity model perturbation $\delta\mathbf{c}$ with the data perturbation $\mathbf{G}_c\delta\mathbf{c}$

$$\mathbf{f}(\mathbf{c}_k + \delta\mathbf{c}) = \mathbf{f}(\mathbf{c}_k) + \mathbf{G}_c\delta\mathbf{c} + \text{higher order terms.} \tag{3.12}$$

The calculation of this linear operator is given in Chapter 2 (from equation (2.27) to equation (2.31). So, using equation (2.31), I have

$$(\mathbf{G}_c\delta\mathbf{c})(x_r,t;x_s) = \int_v dv(x) g(x_r,t;x,0) * \frac{\partial^2 p(x,t;x_s)}{\partial t^2}\frac{2\delta c(x)}{c(x)^3}, \tag{3.13}$$

with $g(x,t;x',t')$ the Green's function and $p(x,t;x_s)$ the pressure wavefield, both associated with the velocity model $c(x)$.

By definition the transpose operator $\mathbf{G}_c^t$ obeys

$$< \delta\hat{\mathbf{p}}, \mathbf{G}_c\delta\mathbf{c} > = < \mathbf{G}_c^t\delta\hat{\mathbf{p}}, \delta\mathbf{c} > \qquad \text{for any } \delta\hat{\mathbf{p}} \text{ and } \delta\mathbf{c}, \tag{3.14}$$

and the norm in the space domain is given by

$$\begin{aligned} < \delta\hat{\mathbf{c}}, \delta\mathbf{c} > &= \delta\hat{\mathbf{c}}^t\delta\mathbf{c} = \delta\mathbf{c}^t\delta\hat{\mathbf{c}} \\ &= \int_v dv(x)\delta\hat{c}(x)\delta c(x). \end{aligned} \tag{3.15}$$

Therefore, using equations (3.15) and (3.4), equation (3.14) becomes

$$\sum_s \int dt \sum_r \delta\hat{p}(x_r,t;x_s)(\mathbf{G}_c\delta\mathbf{c})(x_r,t;x_s) = \int_v dv(x)(\mathbf{G}_c^t\delta\hat{\mathbf{p}})(x)\delta c(x), \tag{3.16}$$

and, using equation (3.13)

$$\begin{aligned} \sum_s \int dt \sum_r \delta\hat{p}(x_r,t;x_s) \int_v dv(x) g(x_r,t;x,0) * \frac{\partial^2 p(x,t;x_s)}{\partial t^2}\frac{2\delta c(x)}{c(x)^3} \\ = \int_v dv(x)(\mathbf{G}_c^t\delta\hat{\mathbf{p}})(x)\delta c(x). \end{aligned} \tag{3.17}$$

Rewriting equation (3.17), I obtain

$$\int_v dv(x)\delta c(x)\left[(\mathbf{G}_c^t\delta\hat{\mathbf{p}})(x) - \frac{2}{c(x)^3}\sum_s\int dt\sum_r g(x_r,t;x,0) * \frac{\partial^2 p(x,t;x_s)}{\partial t^2}\delta\hat{p}(x_r,t;x_s)\right] = 0. \quad (3.18)$$

As this equation is valid for any $\delta c(x)$, I get

$$(\mathbf{G}_c^t\delta\hat{\mathbf{p}})(x) = \frac{2}{c(x)^3}\sum_s\int dt\sum_r g(x_r,t;x,0) * \frac{\partial^2 p(x,t;x_s)}{\partial t^2}\delta\hat{p}(x_r,t;x_s), \quad (3.19)$$

and using equation (2.33) results in

$$(\mathbf{G}_c^t\delta\hat{\mathbf{p}})(x) = -\frac{2}{c(x)^3}\sum_s\int dt\sum_r \dot{g}(x_r,t;x,0) * \dot{p}(x,t;x_s)\delta\hat{p}(x_r,t;x_s). \quad (3.20)$$

Let's now define a pressure field $q_2(x,t;x_s)$ that satisfies

$$\frac{1}{c(x)^2}\frac{\partial^2 q_2(x,t;x_s)}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla q_2(x,t;x_s)\right] = \Phi(x,t;x_s); \quad (3.21)$$
$$q_2(x,T;x_s) = 0 \qquad \dot{q}_2(x,T;x_s) = 0,$$

with

$$\Phi(x,t;x_s) = \sum_r \delta(x-x_r)\delta\hat{p}(x_r,t;x_s). \quad (3.22)$$

In terms of the Green's function, the solution of equation (3.21) (back in time) is

$$\begin{aligned} q_2(x,t;x_s) &= \int_v dv(x')g(x,-t;x',0) * \sum_r \delta(x'-x_r)\delta\hat{p}(x_r,t;x_s) \\ &= \sum_r\int_v dv(x')\delta(x'-x_r)\int dt' g(x,0;x',t-t')\delta\hat{p}(x_r,t';x_s) \\ &= \sum_r\int dt' g(x,0;x_r,t-t')\delta\hat{p}(x_r,t';x_s). \end{aligned} \quad (3.23)$$

The derivative with respect to time $\dot{q}_2(x,t;x_s)$ is given by

$$\begin{aligned}\dot{q}_2(x,t;x_s) &= \frac{\partial}{\partial t}\sum_r \int dt' g(x,0;x_r,t-t')\delta\hat{p}(x_r,t';x_s) \\ &= \frac{\partial}{\partial t}\sum_r \int dt' g(x,t'-t;x_r,0)\delta\hat{p}(x_r,t';x_s) \\ &= -\sum_r \int dt' \dot{g}(x,t'-t;x_r,0)\delta\hat{p}(x_r,t';x_s). \end{aligned} \tag{3.24}$$

Returning to equation (3.20) and rewriting it, I have

$$\begin{aligned}&(\mathbf{G}_c^t\delta\hat{\mathbf{p}})(x) \\ &\quad = -\frac{2}{c(x)^3}\sum_s \int dt \sum_r \int dt' \dot{g}(x,t-t';x_r,0)\dot{p}(x,t';x_s)\delta\hat{p}(x_r,t;x_s). \end{aligned} \tag{3.25}$$

Introducing equation (3.24) into equation (3.25), results in

$$\begin{aligned}\hat{\gamma}_2 &= (\mathbf{G}_c^t\delta\hat{\mathbf{p}})(x) \\ &= \frac{2}{c(x)^3}\sum_s \int dt \dot{p}(x,t;x_s)\dot{q}_2(x,t;x_s). \end{aligned} \tag{3.26}$$

The misfit gradient $\hat{\gamma}_2$ is then calculated as the zero-lag correlation between the time derivative of the pressure field $p(x,t;x_s)$ generated using the current velocity model, and the time derivative of the pressure field $q_2(x,t;x_s)$ computed by reverse-time propagation of the seismogram residual $\delta p(x_r,t;x_s)$. The last one represents the diffracted field missing in the current data.

As shown by Luo and Schuster (1991), the wave-equation traveltime inversion reduces to the full waveform inversion when the current velocity model is close to the actual model. This means that the calculated seismogram has similar amplitude to the observed seismogram and they differ only by a small time shift. Figure 3.1 depicts this fact by comparing the functions that are back-propagated in both cases. The waveforms in the figure are hypothetical functions and do not represent the propagation in any medium. The back-propagated fields are similar

except for a constant factor of $2/E$,

$$\delta\tau(x_r, t; x_s) = \frac{2}{E}\delta p(x_r, t; x_s) \tag{3.27}$$

where E is defined by equation (2.5).

3.2 Numerical examples and discussion

A standard computer code for full waveform inversion was provided by the University of Utah Consortium, and I implemented the parallelised version of this code on the Cray T3D in the Edinburgh Parallel Computing Centre (EPCC). A finite difference scheme was codified for the forward and reverse-time propagation.

As shown in Figure 2.5, I applied the full waveform inversion to a model with crosshole configuration. The transmitted arrivals were muted out prior to back-propagation and the initial velocity model corresponded to the output from the standard traveltime inversion. After 4 iterations, it succeeded in reconstructing the high frequency details of the synthetic fault model.

A second model consists of two dipping interfaces, the first one with 30 degree slope and the deeper one with a gentle slope of 9 degrees, see Figure 3.2. A set of 15 shots and 31 receivers, evenly spaced, is located at the surface position. The velocity model is digitised with a regular grid of 62x141 points. The sampling interval for the grid is 1.5 m. The sample rate is 0.0002 s and 1000 samples in time are calculated for each trace. The source function is a Ricker wavelet with 60 Hz as a peak frequency. Again, the density values are calculated using the Gardner's formula (Gardner et al., 1974; Zhou, 1995; Zhou et al., 1995), equation (2.8). As in the traveltime inversion method a preconditioning function is needed in order to accelerate the convergence.

The starting model is a constant velocity field of 2000 m/s and the first arrivals are muted out. The step length is kept constant. Figure 3.3 shows the full waveform inversion after 1, 2, 3 and 4 iterations. The corresponding seismograms for shot 1 are shown in Figure 3.4. Figure 3.5 depicts the behaviour of the residuals.

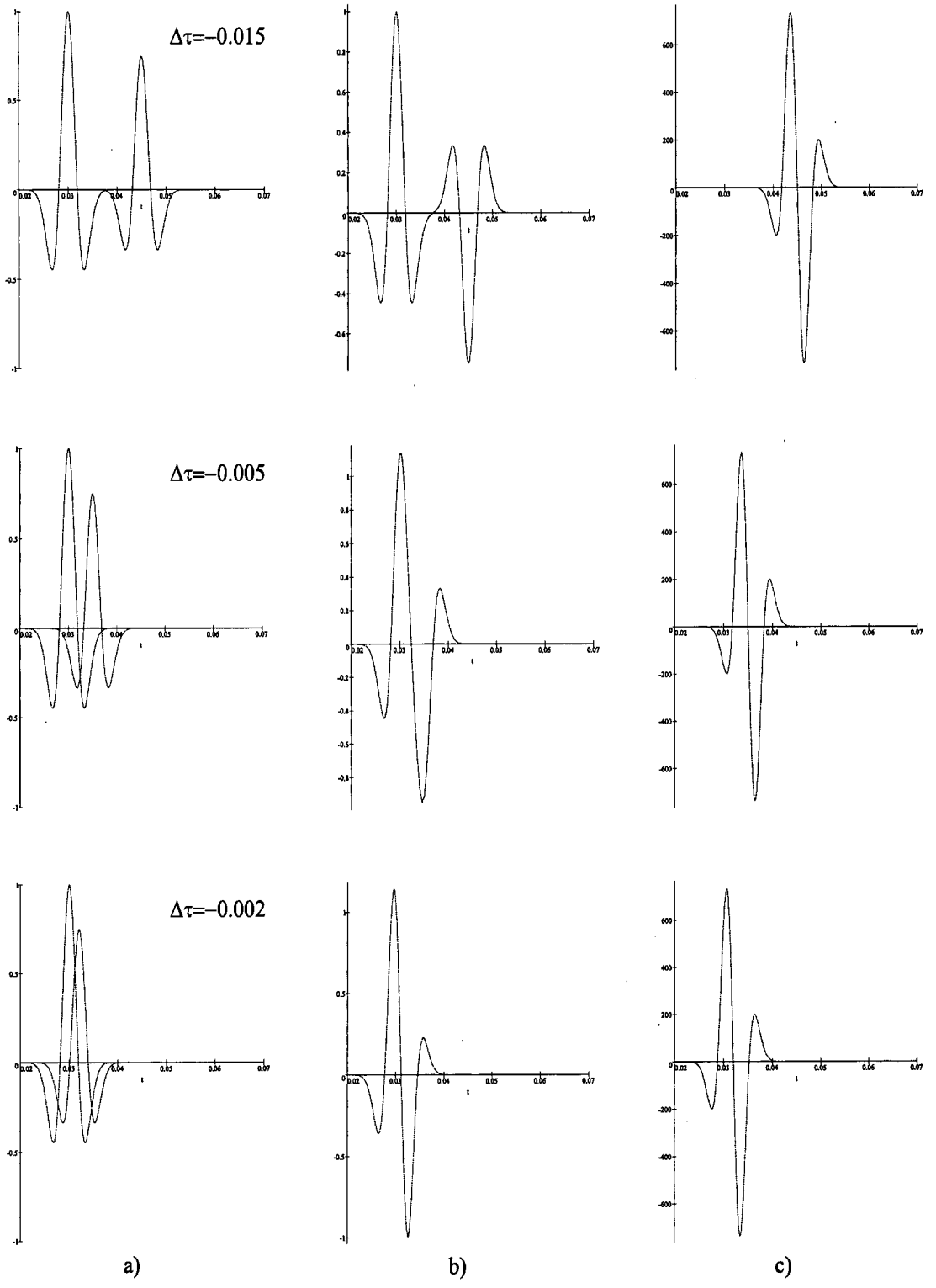


Figure 3.1: Traveltime inversion reduces to full waveform inversion when $\Delta\tau$ is small. a) $p(t)_{obs}$ and $p(t)_{cal}$, b) $p(t)_{obs} - p(t)_{cal}$, c) $\dot{p}(t + \Delta\tau)_{obs}$.

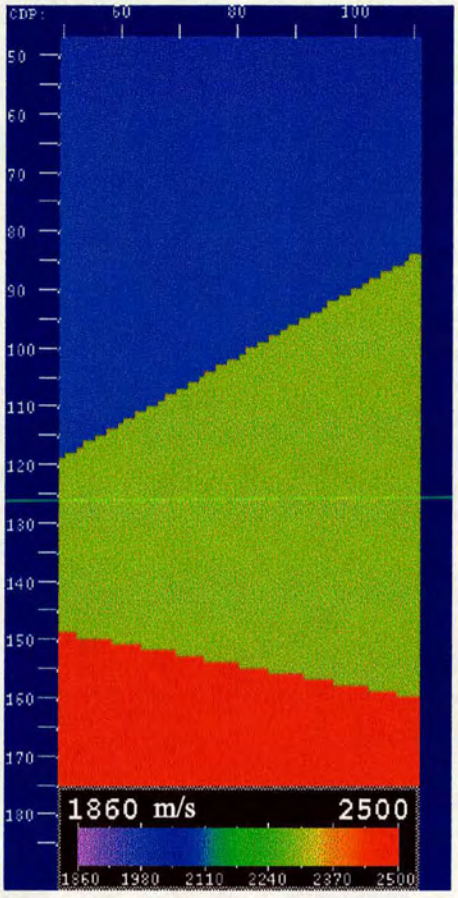


Figure 3.2: Velocity model with dipping interfaces.

The oscillation in the residual curve is a consequence of keeping the step length fixed. The inversion is converging to a minimum, but the step length is so big that it does not allow small changes in the gradient. Therefore, the velocity model oscillates (see Figure 3.3) in the same way as the residual does. The results from iterations 1 and 3 look alike. Also, the results from iterations 2 and 4 are similar. This effect can also be seen on the shot gathers.

Artifacts with a regular pattern appear after the first iteration and are related to the peak frequency of the Ricker wavelet. Figure 3.6 shows the inversion after 5 iterations using two different frequencies for the Ricker wavelet. As the peak frequency increases, the number of artifacts increases as well. The velocity values increase and decrease regularly following the artifact pattern. This corresponds

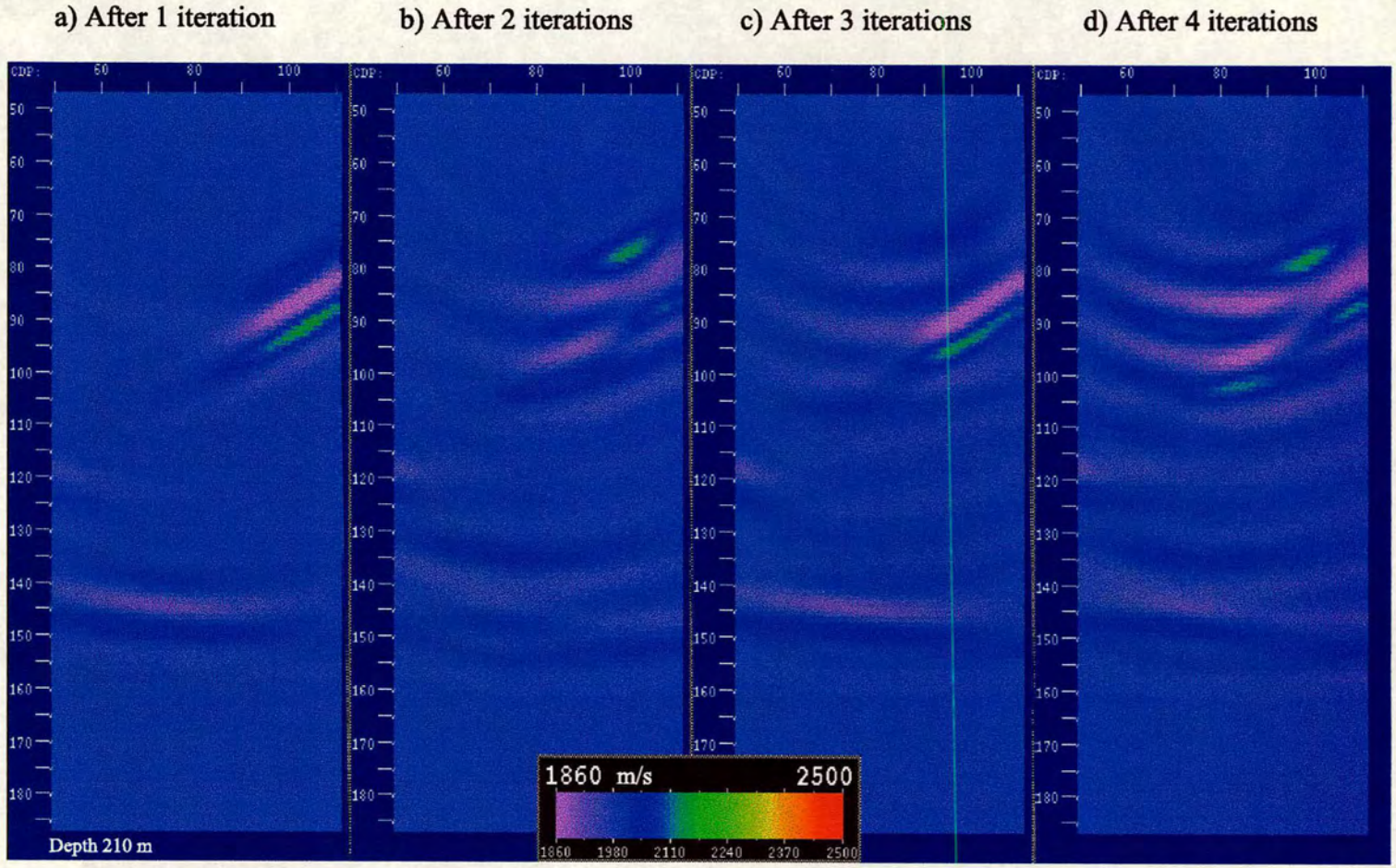


Figure 3.3: Full waveform inversion of the model with dipping interfaces.

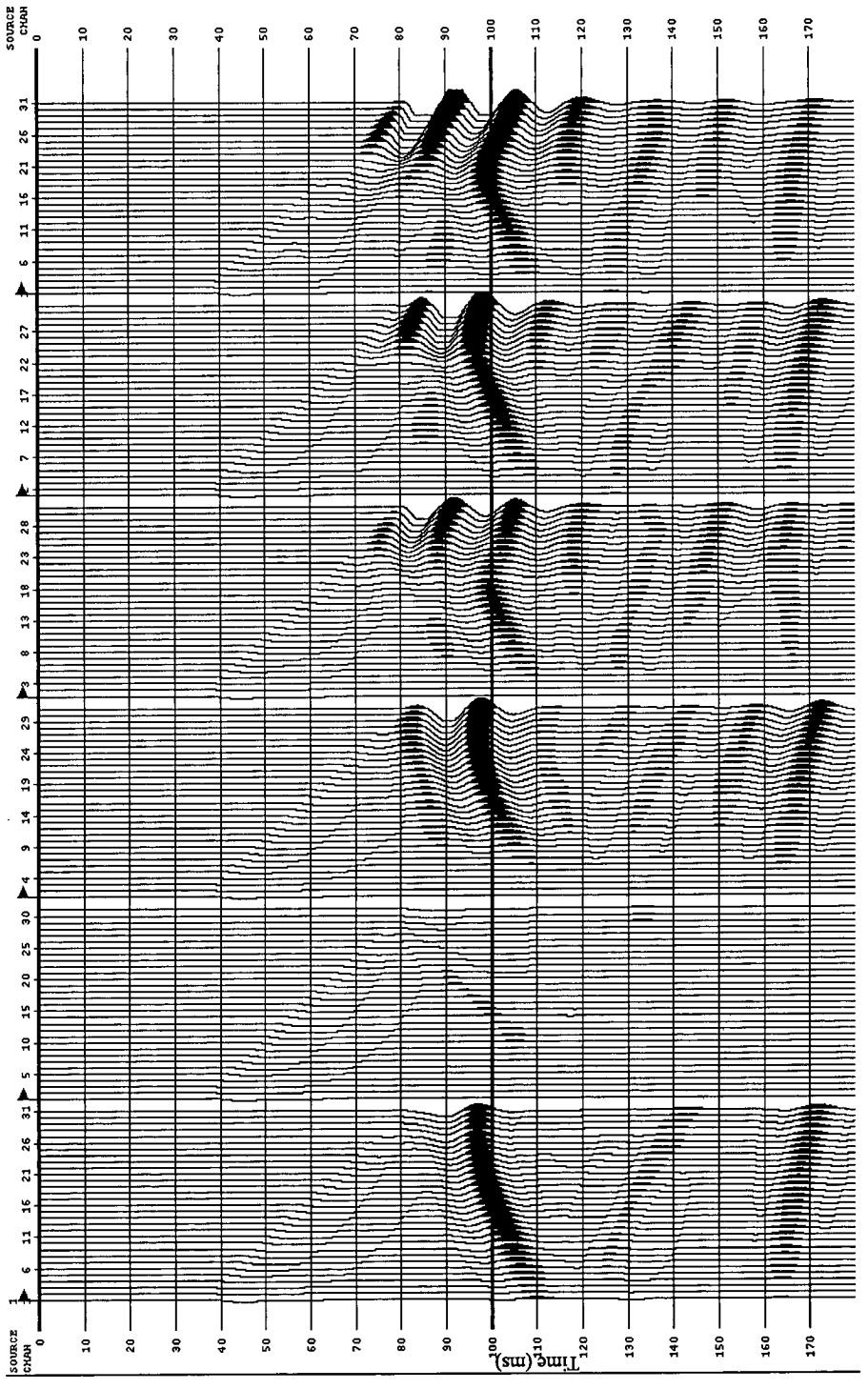


Figure 3.4: Shot gathers from the full waveform inversion.

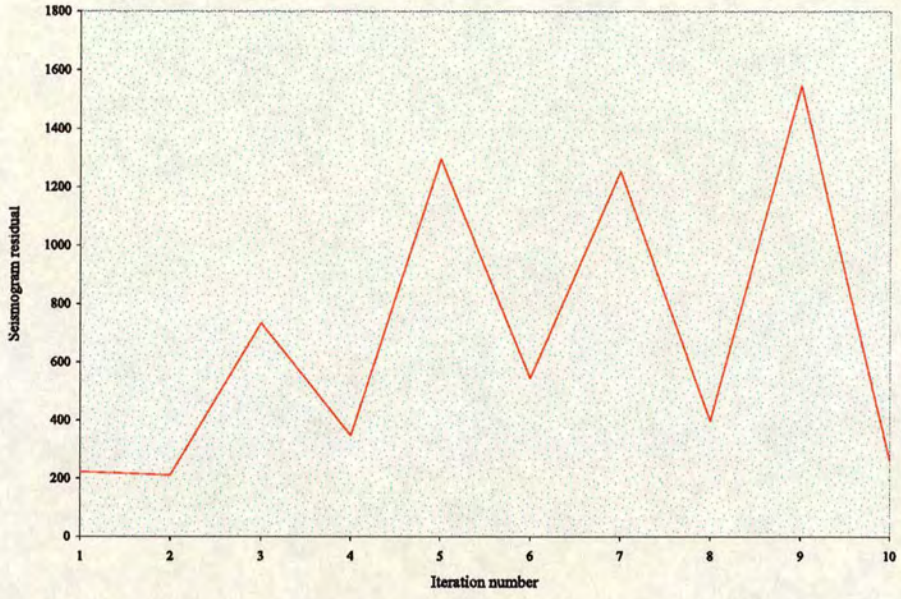


Figure 3.5: Residuals of the dipping layers inversion.

to positive and negative values of the gradient generated by periodic changes from peaks to troughs in the residual seismogram.

Figure 3.7 depicts a different velocity model with the purpose of illustrating the behaviour of the artifacts when the mute of the first arrivals is not applied, when it is applied, and when the normalization of the seismograms takes place after the mute. For this example 5 iterations of full waveform inversion were executed in each case. If the complete seismograms are considered, the high amplitude of the first arrivals dominates the inversion. The artifacts seem not to be present but the inversion is constrained to the upper part of the model, where the direct arrivals provide only near-surface information.

In contrast, in crosshole configurations the direct arrivals cross the model from one side to the other, carrying a complete set of information concerning the distribution of the velocity field. This is one of the most important factors to be considered when surveys with different geometry configurations are going to be inverted. Figures 3.7c and 3.7d show why the first arrivals should be muted out

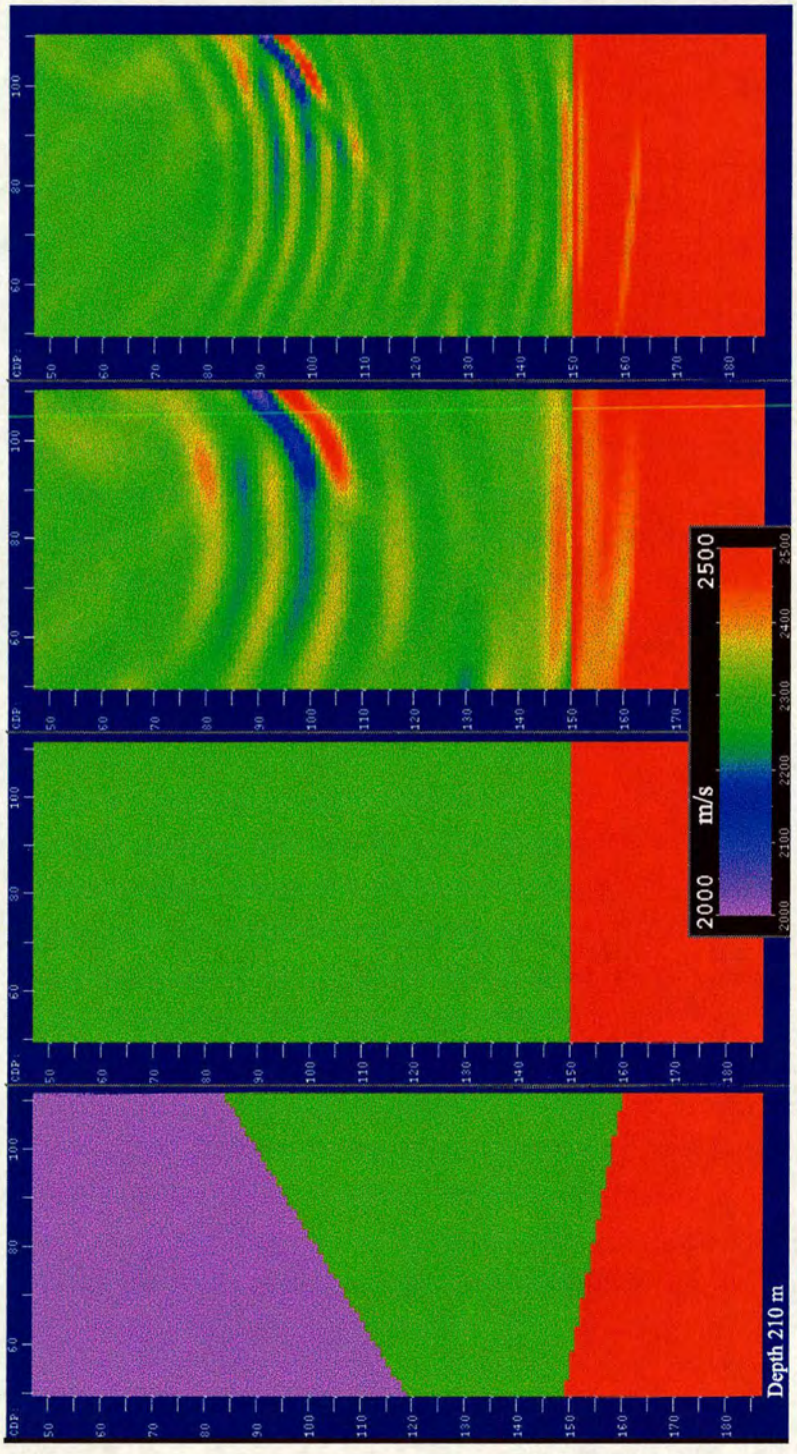


Figure 3.6: Full waveform inversion using two different Ricker wavelets.

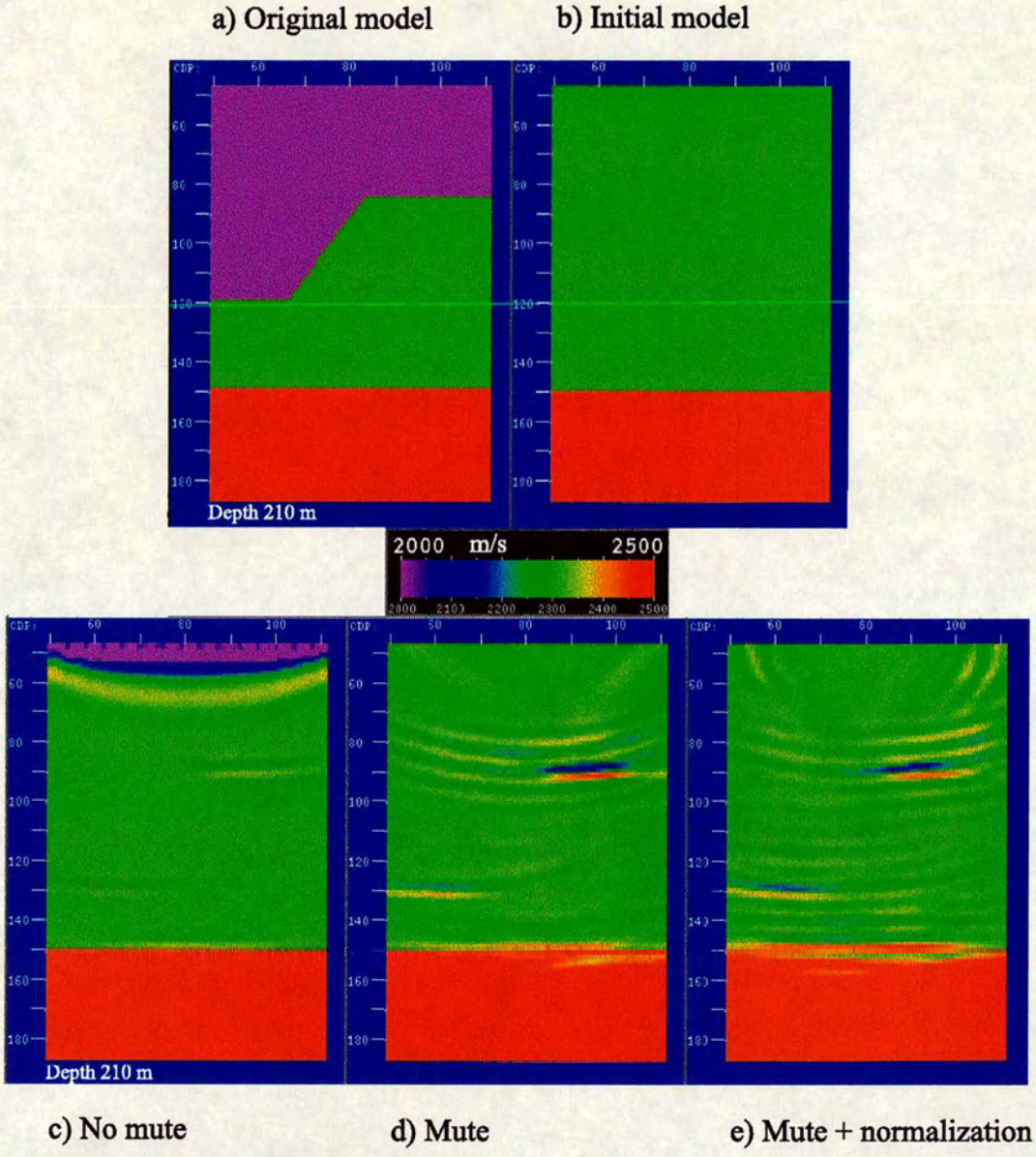


Figure 3.7: Different factors that affect the full waveform inversion.

from the seismograms when inverting datasets with surface geometry configuration using the full waveform inversion method. After muting out the first arrivals, the reflection events start appearing. The same criterion applies to the standard traveltime inversion and to the traveltime-offset inversion method.

Figure 3.7e shows the result when the traces in both seismograms, the observed and the calculated, are normalized to its maximum amplitude after muting out the first arrivals. This process brings up the value of the amplitudes of the remaining waveforms, but keeps the relative amplitude among them. Then, with higher amplitudes the artifacts are more evident than before.

In Figures 3.8 and 3.9, the gradients γ_2 corresponding to a single shot and a single receiver are plotted for a crosshole configuration and a surface reflection geometry, respectively. The original model now is a three flat layer medium with velocities 2000 m/s, 2300 m/s, and 2500 m/s, and interfaces at 109.5 m and 154.5 m depth (Figure 3.10a). The initial velocity model keeps the depth positions and the velocities are 10% slower than the original ones.

In the crosshole configuration, the direct arrival dominates the gradient and the absolute value of this gradient is one order of magnitude bigger than in the surface geometry. In the surface reflection configuration, the mute is not removing the first arrival signal completely. The remaining part of this signal generates the first artifact, located in depth between the nodes 60 and 70. The deeper artifacts seen in Figure 3.9 are also present in Figure 3.8, but they are negligible compared to the major contribution of the direct arrival to the gradient. As pointed out by Rowbotham and Goulty (1993), an effective removal of the direct waves is essential for optimising the quality of the image.

The same model with surface reflection geometry is used to find out if the initial velocity model changes the behaviour of the artifacts. The results are shown in Figure 3.10, where 2000 m/s and 2500 m/s are the values of the constant velocity fields used as starting models. The artifacts follow the same regular pattern and, again, the first artifact corresponds to the remaining direct arrival not muted out

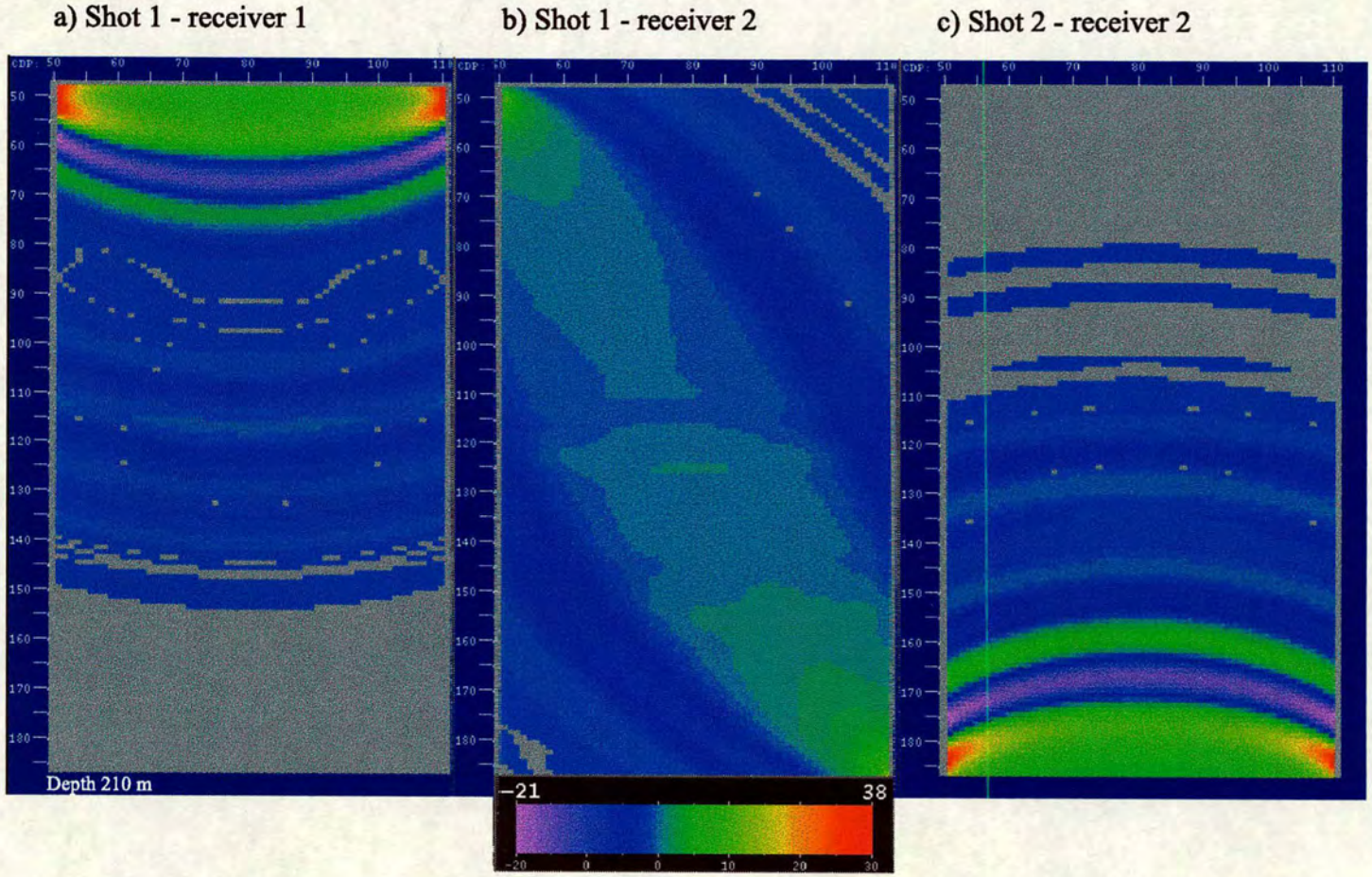


Figure 3.8: Gradient for a crosshole configuration. Shot 1 coordinates (50,50), shot 2 coordinates (50,186), receiver 1 coordinates (110,50), receiver 2 coordinates (110,186).

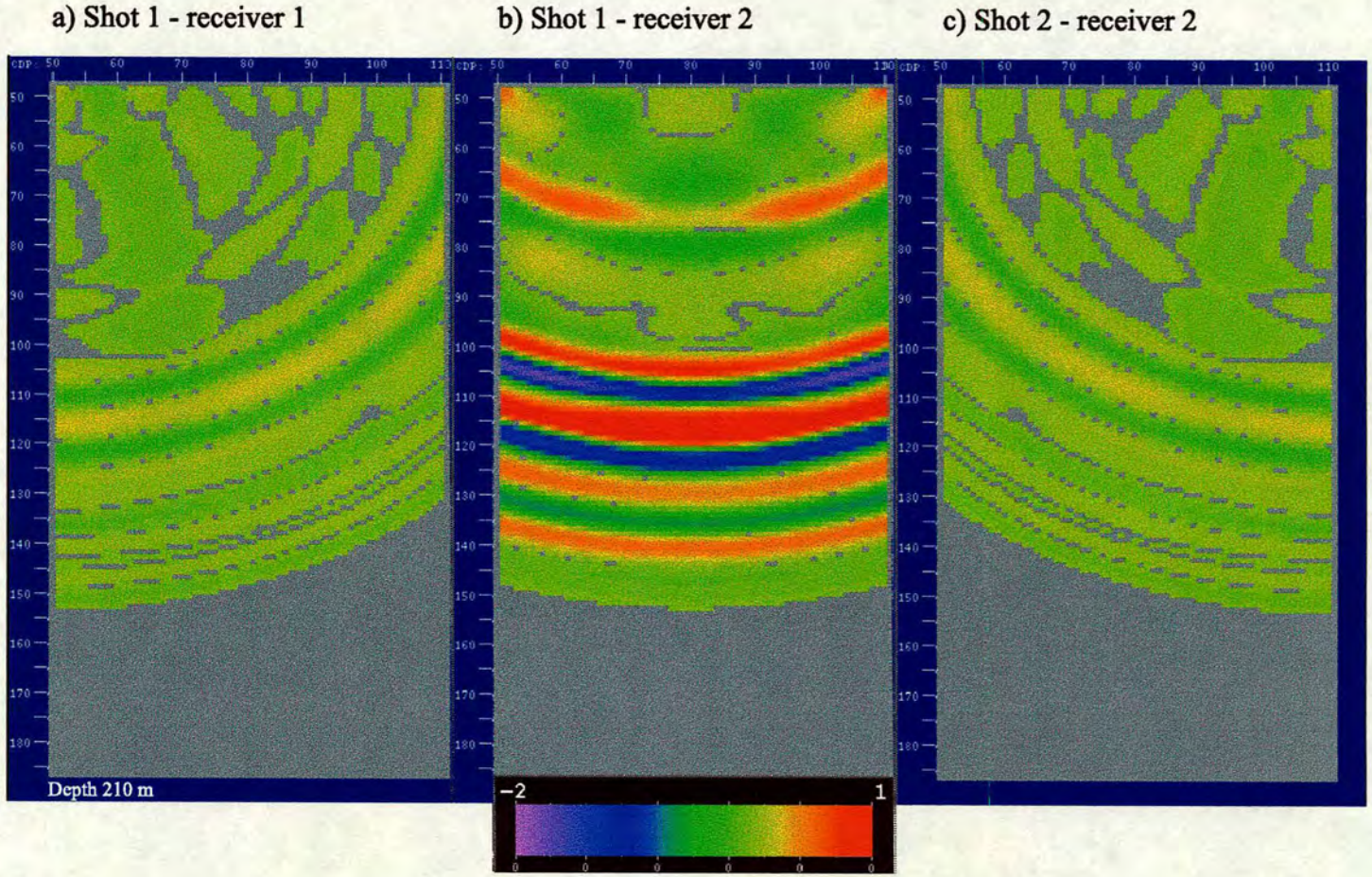


Figure 3.9: Gradient for a surface reflection geometry. Shot 1 coordinates (50,47), shot 2 coordinates (111,47), receiver 1 coordinates (51,47), receiver 2 coordinates (110,47).

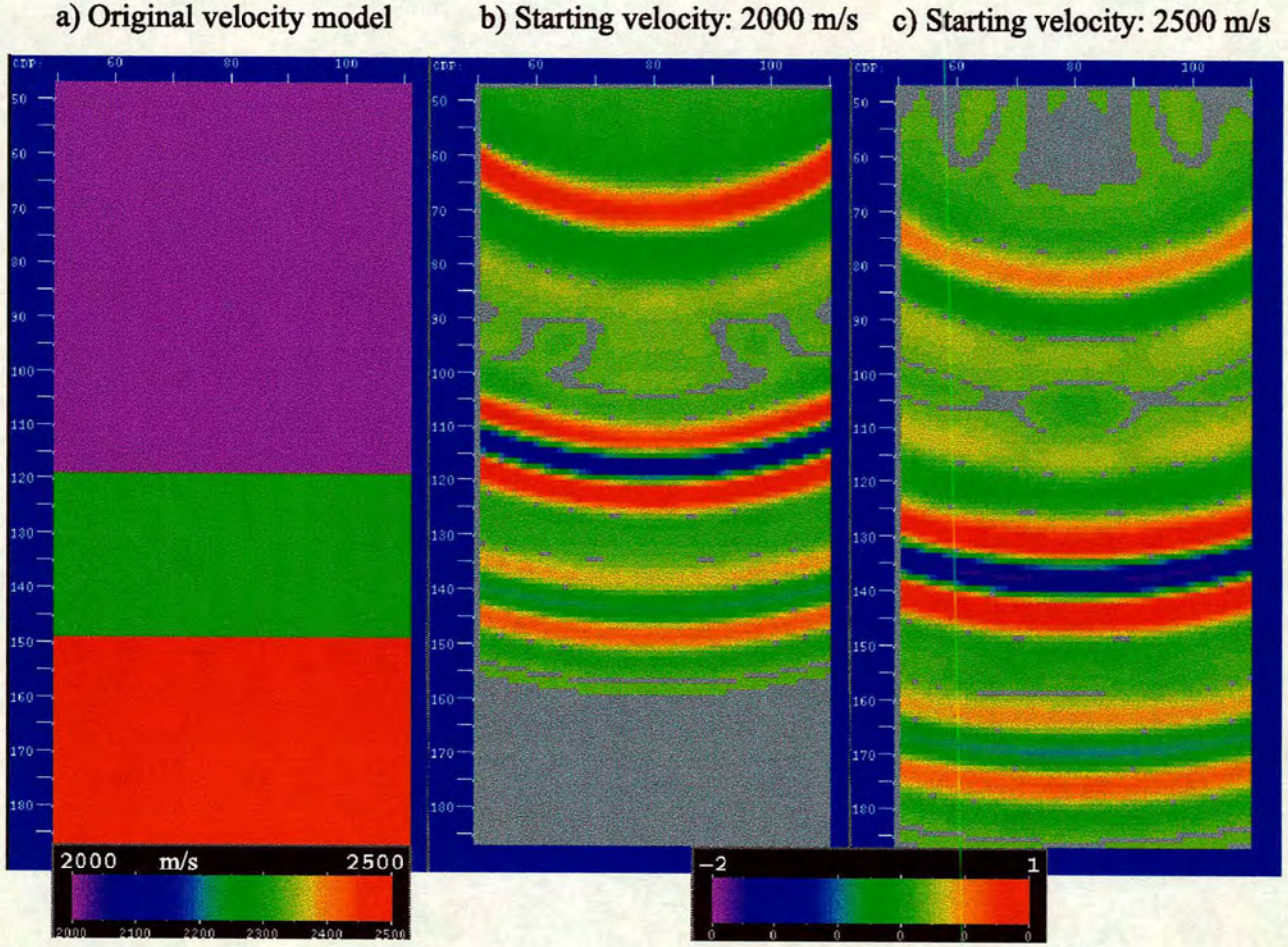


Figure 3.10: Gradient of the full waveform inversion using two different starting velocity models. Shot coordinates (50,47), receiver coordinates (110,47).

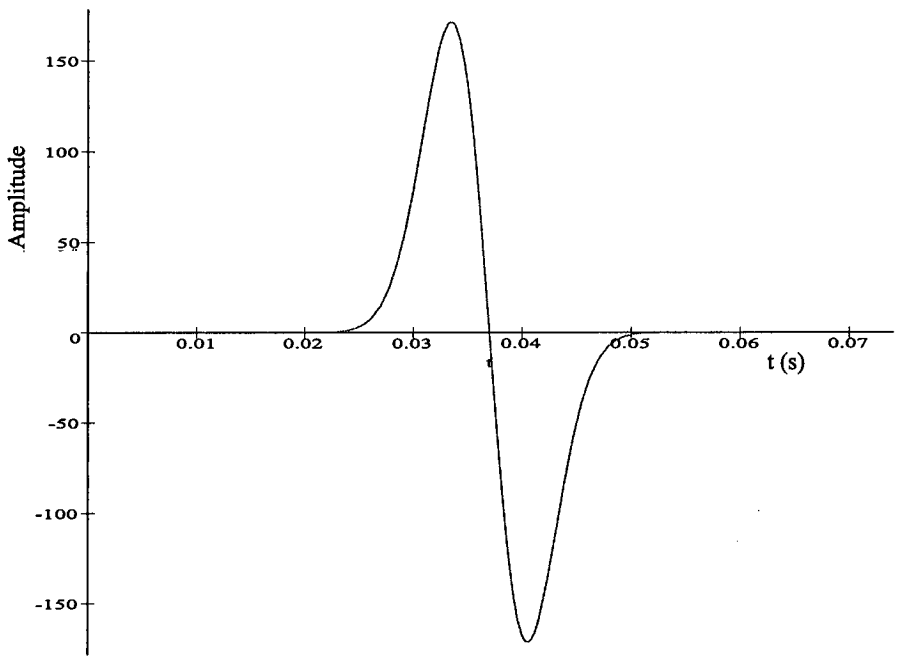


Figure 3.11: Gaussian-derivative wavelet (120 Hz).

correctly.

Next the model in Figure 3.2 (surface reflection configuration) is again used, but with the derivative of a Gaussian function (Savic, 1995) as the source function instead of the Ricker wavelet. The new wavelet, Figure 3.11 is given by the following equation:

$$s(t) = -2\xi t \exp(-\xi t^2) \tag{3.28}$$

where $\xi = 4 \ln 2 \nu^2$ and ν is the peak frequency of the wavelet.

Figure 3.12 depicts the full waveform inversion using the new wavelet after 10, 20, 30, and 40 iterations. The mute was not applied but the initial velocity model corresponded to the velocity of the first layer (2000 m/s) and the first 7 cells in depth were forced to keep the correct velocities. The behaviour of the inversion is completely different from the one with the Ricker wavelet; Figure 3.13 displays the comparison. There are no artifacts present and the interfaces show a

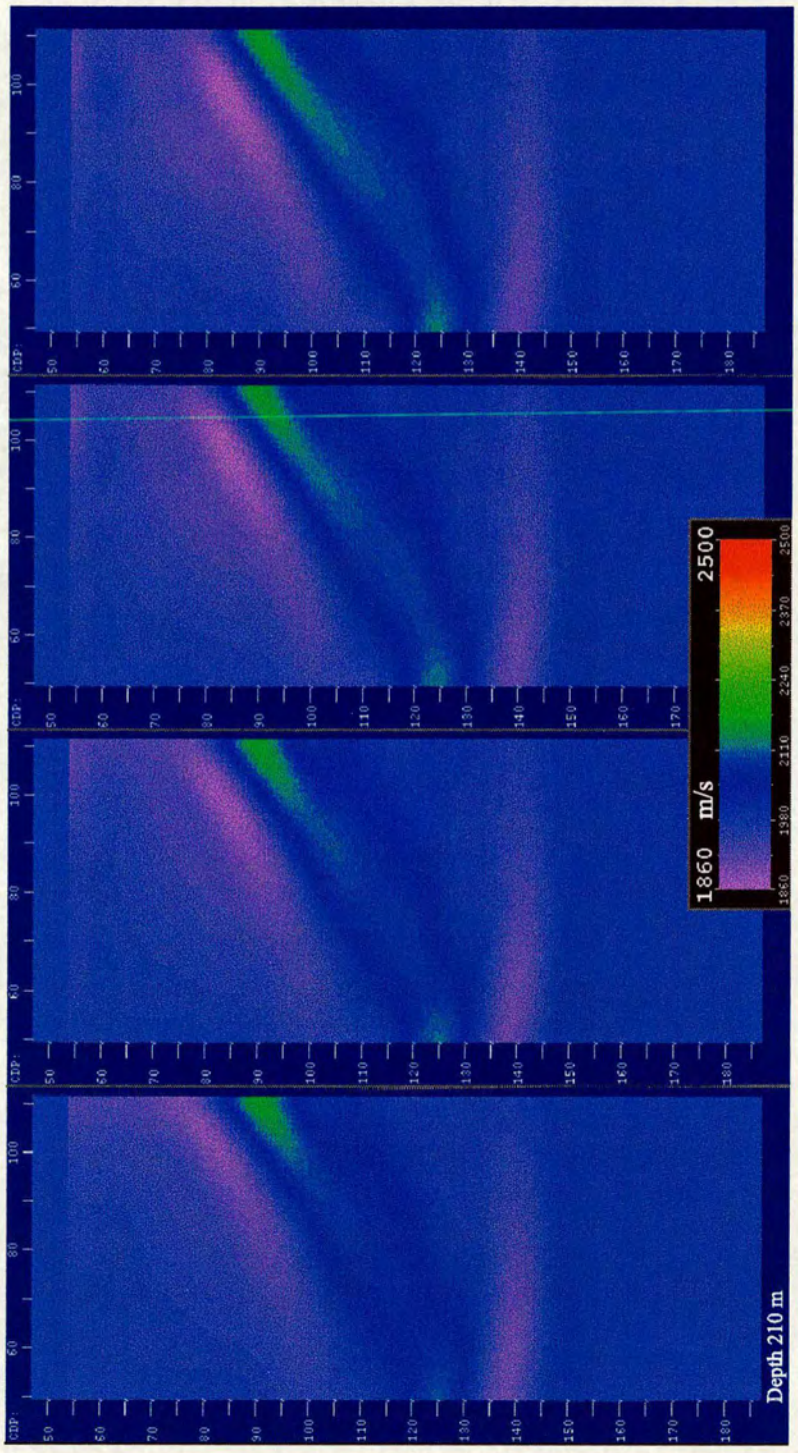
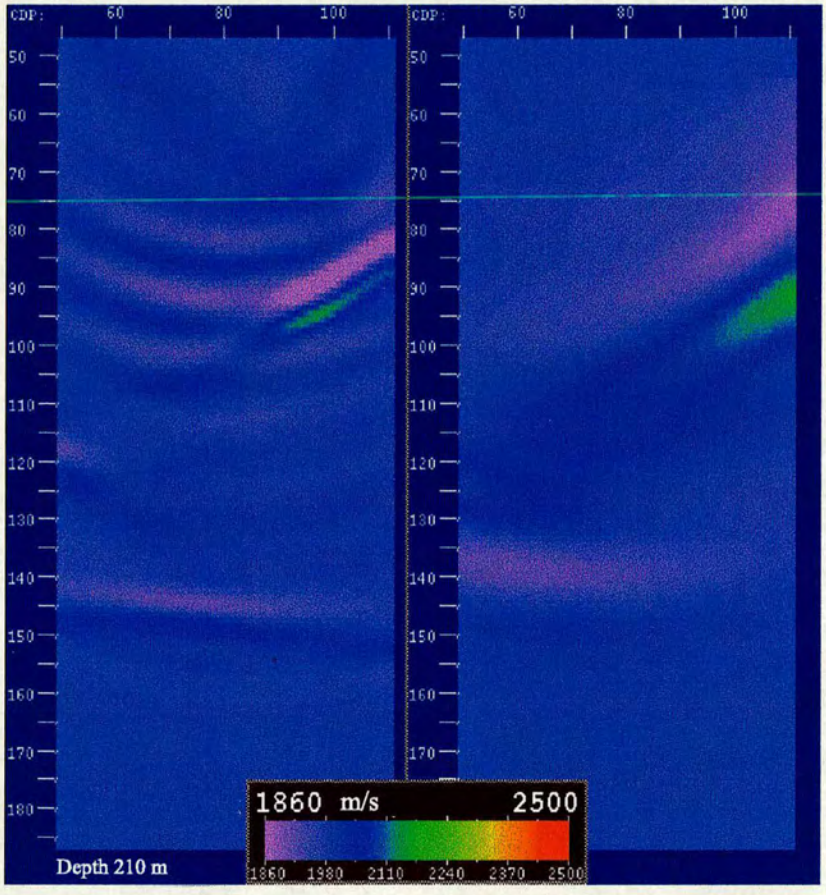


Figure 3.12: Full waveform inversion using the derivative of a Gaussian function as the source function.



a) Ricker wavelet **b) Derivative of a Gaussian function wavelet**

Figure 3.13: Full waveform inversion using two different wavelets.

better resolution. However, the result is still constrained by the minimum velocity allowed during the inversion, 1800 m/s for this example.

3.3 Summary

The full waveform inversion back-propagates the differences between the observed and the calculated seismograms, considering these differences as missing diffracted fields. This method recovers the fine details present in the original model when the starting velocity model contains the low frequency features of the original model. It succeeds in achieving the inversion of a synthetic fault model with crosshole configuration, but some artifacts are present if surface reflection datasets with a Ricker wavelet are inverted. The artifacts are related to the frequency spectrum of the source wavelet and to the amplitude of the different waveforms. Also, an inappropiate mute of the first arrivals can introduce additional artifacts. In Chapter 6, I combine the full waveform inversion with the new traveltime-offset inversion method and some further examples are considered in order to find out an explanation for the behaviour of the full waveform inversion.

Chapter 4

Wave-equation traveltime-offset inversion

A velocity estimation approach is set up using a bi-dimensional cross-correlation function. The new method modifies the standard wave-equation traveltime inversion (Luo and Schuster, 1991) by introducing the cross-correlation function in the time and space domains simultaneously. The velocity model is updated using the current velocity model, the observed pressure field and the field computed by reverse propagation of two pseudo-residual functions acting as sources in particular locations. The shot gathers are moved as a whole; that is, the traces are not shifted independently in the common shot gathers, and then the moveout of the real shot gathers is maintained during the back-propagation.

In this chapter, I present the theory involved in the proposed wave-equation traveltime-offset inversion method. I define a new bi-dimensional connective function and a new traveltime-offset misfit function. Finally, using the steepest descent method I derive an expression for the gradient function that updates the velocity field at each iteration.

4.1 Theory

As in the standard traveltime inversion, let $p(x_r, t; x_s)_{obs}$ and $p(x_r, t; x_s)_{cal}$ denote, respectively, the observed and calculated pressure seismograms at receiver location x_r due to a line source at the location x_s. The calculated seismogram satisfies the acoustic wave equation:

$$\frac{1}{c(x)^2}\frac{\partial^2 p(x,t;x_s)}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla p(x,t;x_s)\right] = s(t;x_s), \tag{4.1}$$

where $s(t;x_s)$, $\rho(x)$, $c(x)$ correspond to the source function, the density function and the wave speed, respectively.

The bi-dimensional cross-correlation function is defined as a function of a offset shift σ and a time shift τ by

$$f(\sigma,\tau;x_s) = \int dt \int dx_r \frac{p(x_r+\sigma,t+\tau;x_s)_{obs}}{A(x_r;x_s)_{obs}} p(x_r,t;x_s)_{cal}, \tag{4.2}$$

where $A(x_r;x_s)_{obs}$ is the maximum amplitude of the observed seismogram. Instead of moving a real trace over a calculated trace, the bi-dimensional function moves the full observed shot gather over the full calculated shot gather. The best match is reached when the function f has a maximum

$$f(\Delta\sigma,\Delta\tau;x_s) = \max\{f(\sigma,\tau;x_s)\}\,|\tau\in[-T,T],\sigma\in[-X,X]\}. \tag{4.3}$$

The derivatives in both dimensions should be zero and the following equations have to be satisfied simultaneously

$$\begin{aligned} \left[\frac{\partial f(\sigma,\tau;x_s)}{\partial \tau}\right]_{\sigma=\Delta\sigma,\tau=\Delta\tau} &= \dot{f}=0 \\ \left[\frac{\partial f(\sigma,\tau;x_s)}{\partial \sigma}\right]_{\sigma=\Delta\sigma,\tau=\Delta\tau} &= f'=0. \end{aligned} \tag{4.4}$$

Then,

$$\begin{aligned} \dot{f} &= \int dt \int dx_r \frac{1}{A(x_r;x_s)_{obs}} \dot{p}(x_r+\Delta\sigma,t+\Delta\tau;x_s)_{obs} p(x_r,t;x_s)_{cal} = 0 \\ f' &= \int dt \int dx_r \frac{1}{A(x_r;x_s)_{obs}} p'(x_r+\Delta\sigma,t+\Delta\tau;x_s)_{obs} p(x_r,t;x_s)_{cal} = 0. \end{aligned} \tag{4.5}$$

The misfit function S, defined as

$$S = \frac{1}{2}\sum_s [\Delta\sigma(x_s)^2 + \beta\Delta\tau(x_s)^2], \tag{4.6}$$

has a minimum when the predicted shot gather best matches the observed shot gather, i.e. when the cross-correlation function has a maximum. I introduce the parameter β to keep the dimensions of this equation consistent.

For simplicity I use the steepest descent method in order to find the velocity model $c(x)$ that minimises S. Hence, the new velocity field is calculated using the following equation:

$$c(x)_{k+1} = c(x)_k + \alpha_k \gamma(x)_k, \tag{4.7}$$

where $\gamma(x)$ is the steepest descent direction of the misfit function S, given by

$$\gamma(x) = -\frac{\partial S}{\partial c(x)} = -\sum_s \frac{\partial \Delta\sigma}{\partial c(x)} \Delta\sigma(x_s) - \beta \sum_s \frac{\partial \Delta\tau}{\partial c(x)} \Delta\tau(x_s) = \gamma_\sigma(x) + \beta\gamma_\tau(x), \tag{4.8}$$

and α_k is a constant scaling factor, either analytically estimated (Tarantola, 1986; Tarantola, 1987) or chosen by trial and error. Using the rule for implicit differentiation

$$\begin{aligned} \frac{\partial \dot{f}}{\partial c(x)} &= \frac{\partial \dot{f}}{\partial \Delta\tau} \frac{\partial \Delta\tau}{\partial c(x)} + \frac{\partial \dot{f}}{\partial \Delta\sigma} \frac{\partial \Delta\sigma}{\partial c(x)} \\ \frac{\partial f'}{\partial c(x)} &= \frac{\partial f'}{\partial \Delta\tau} \frac{\partial \Delta\tau}{\partial c(x)} + \frac{\partial f'}{\partial \Delta\sigma} \frac{\partial \Delta\sigma}{\partial c(x)}, \end{aligned} \tag{4.9}$$

and solving the system of equations simultaneously, I get

$$\frac{\partial \Delta\sigma}{\partial c(x)} = \frac{\dfrac{\partial f'}{\partial c(x)} - \dfrac{\partial \dot{f}}{\partial c(x)} \dfrac{\partial f'/\partial \Delta\tau}{\partial \dot{f}/\partial \Delta\tau}}{\dfrac{\partial f'}{\partial \Delta\sigma} - \dfrac{\partial \dot{f}}{\partial \Delta\sigma} \dfrac{\partial f'/\partial \Delta\tau}{\partial \dot{f}/\partial \Delta\tau}} \tag{4.10}$$

$$\frac{\partial \Delta\tau}{\partial c(x)} = \frac{\dfrac{\partial \dot{f}}{\partial c(x)} - \dfrac{\partial f'}{\partial c(x)} \dfrac{\partial \dot{f}/\partial \Delta\sigma}{\partial f'/\partial \Delta\sigma}}{\dfrac{\partial \dot{f}}{\partial \Delta\tau} - \dfrac{\partial f'}{\partial \Delta\tau} \dfrac{\partial \dot{f}/\partial \Delta\sigma}{\partial f'/\partial \Delta\sigma}}. \tag{4.11}$$

Using the connective functions, equation (4.5), and assuming that mixed second order partial derivatives commute, I have

$$\begin{aligned}
\frac{\partial \dot{f}}{\partial c(x)} &= \int dt \int dx_r \frac{1}{A(x_r;x_s)_{obs}} \dot{p}(x_r+\Delta\sigma, t+\Delta\tau; x_s)_{obs} \frac{\partial p(x_r,t;x_s)_{cal}}{\partial c(x)} \\
\frac{\partial \dot{f}}{\partial \Delta\tau} &= \int dt \int dx_r \frac{1}{A(x_r;x_s)_{obs}} \ddot{p}(x_r+\Delta\sigma, t+\Delta\tau; x_s)_{obs} p(x_r,t;x_s)_{cal} \\
\frac{\partial \dot{f}}{\partial \Delta\sigma} &= \int dt \int dx_r \frac{1}{A(x_r;x_s)_{obs}} \dot{p}'(x_r+\Delta\sigma, t+\Delta\tau; x_s)_{obs} p(x_r,t;x_s)_{cal},
\end{aligned} \tag{4.12}$$

$$\begin{aligned}
\frac{\partial f'}{\partial c(x)} &= \int dt \int dx_r \frac{1}{A(x_r;x_s)_{obs}} p'(x_r+\Delta\sigma, t+\Delta\tau; x_s)_{obs} \frac{\partial p(x_r,t;x_s)_{cal}}{\partial c(x)} \\
\frac{\partial f'}{\partial \Delta\tau} &= \int dt \int dx_r \frac{1}{A(x_r;x_s)_{obs}} \dot{p}'(x_r+\Delta\sigma, t+\Delta\tau; x_s)_{obs} p(x_r,t;x_s)_{cal} \\
\frac{\partial f'}{\partial \Delta\sigma} &= \int dt \int dx_r \frac{1}{A(x_r;x_s)_{obs}} p''(x_r+\Delta\sigma, t+\Delta\tau; x_s)_{obs} p(x_r,t;x_s)_{cal}.
\end{aligned} \tag{4.13}$$

Let d_1 and d_2 denote respectively the denominators in equations (4.10) and (4.11) multiplied by $A(x_r;x_s)_{obs}$:

$$\begin{aligned}
d_1 &= A(x_r;x_s)_{obs} \frac{\partial f'}{\partial \Delta\sigma} - A(x_r;x_s)_{obs} \frac{\partial \dot{f}}{\partial \Delta\sigma} \frac{\partial f'/\partial\Delta\tau}{\partial \dot{f}/\partial\Delta\tau} \\
&= \int dt \int dx_r p''_{obs} p_{cal} - \frac{\left[\int dt \int dx_r \ddot{p}'_{obs} p_{cal}\right]^2}{\int dt \int dx_r \ddot{p}_{obs} p_{cal}}
\end{aligned} \tag{4.14}$$

$$\begin{aligned}
d_2 &= A(x_r;x_s)_{obs} \frac{\partial \dot{f}}{\partial \Delta\tau} - A(x_r;x_s)_{obs} \frac{\partial f'}{\partial \Delta\tau} \frac{\partial \dot{f}/\partial\Delta\sigma}{\partial f'/\partial\Delta\sigma} \\
&= \int dt \int dx_r \ddot{p}_{obs} p_{cal} - \frac{\left[\int dt \int dx_r \ddot{p}'_{obs} p_{cal}\right]^2}{\int dt \int dx_r p''_{obs} p_{cal}}.
\end{aligned} \tag{4.15}$$

I excluded the variables of the pressure field functions just for simplicity of pre-

sentation. These definitions, together with eqs. (4.10-4.13) yield:

$$\frac{\partial \Delta\sigma}{\partial c(x)} = \frac{1}{d_1}\left[\int dt \int dx_r p'_{obs}\frac{\partial p_{cal}}{\partial c(x)} - \int dt \int dx_r \dot{p}_{obs}\frac{\partial p_{cal}}{\partial c(x)}\frac{\int dt \int dx_r \ddot{p}'_{obs} p_{cal}}{\int dt \int dx_r \ddot{p}_{obs} p_{cal}}\right] \tag{4.16}$$

$$\frac{\partial \Delta\tau}{\partial c(x)} = \frac{1}{d_2}\left[\int dt \int dx_r \dot{p}_{obs}\frac{\partial p_{cal}}{\partial c(x)} - \int dt \int dx_r p'_{obs}\frac{\partial p_{cal}}{\partial c(x)}\frac{\int dt \int dx_r \ddot{p}'_{obs} p_{cal}}{\int dt \int dx_r p''_{obs} p_{cal}}\right]. \tag{4.17}$$

The derivative of $p(x_r, t; x_s)_{cal}$ with respect to $c(x)$ is:

$$\frac{\partial p(x_r, t; x_s)_{cal}}{\partial c(x)} = -\frac{2}{c(x)^3}\dot{g}(x_r, t; x, 0) * \dot{p}(x, t; x_s), \tag{4.18}$$

where the asterisk represents time convolution (for the calculation of this expression see equations (2.27)-(2.37) in Chapter 2). $p(x, t; x_s)$ satisfies equation (4.1) and the corresponding Green's function $g(x, t; x', t')$ obeys the wave equation:

$$\frac{1}{c(x)^2}\frac{\partial^2 g(x, t;', t')}{\partial t^2} - \rho(x)\nabla.\left[\frac{1}{\rho(x)}\nabla g(x, t; x', t')\right] = \delta(x - x')\delta(t - t') \tag{4.19}$$
$$g(x, t; x', t') = 0, \qquad \dot{g}(x, t; x', t') = 0, \qquad t \leq t'.$$

Using equation 4.18 with the convolutional term expanded, and exchanging the integrals I can write

$$\begin{aligned}
&\int dt \int dx_r p'(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs}\frac{\partial p(x_r, t; x_s)_{cal}}{\partial c(x)} = \\
&\int dt \int dx_r \frac{-2}{c(x)^3} p'(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs}\dot{g}(x_r, t; x, 0) * \dot{p}(x, t; x_s) = \\
&\int dt \int dx_r \frac{-2}{c(x)^3} p'(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs}\int dt' \dot{g}(x_r, t - t'; x, 0)\dot{p}(x, t'; x_s) = \\
&\int dt' \int dx_r \frac{-2}{c(x)^3}\dot{p}(x, t'; x_s)\int dt \dot{g}(x_r, t - t'; x, 0)p'(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs},
\end{aligned} \tag{4.20}$$

and

$$\begin{aligned}
&\int dt \int dx_r \dot{p}(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs} \frac{\partial p(x_r, t; x_s)_{cal}}{\partial c(x)} = \\
&\int dt \int dx_r \frac{-2}{c(x)^3} \dot{p}(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs} \dot{g}(x_r, t; x, 0) * \dot{p}(x, t; x_s) = \\
&\int dt \int dx_r \frac{-2}{c(x)^3} \dot{p}(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs} \int dt' \dot{g}(x_r, t - t'; x, 0) \dot{p}(x, t'; x_s) = \\
&\int dt' \int dx_r \frac{-2}{c(x)^3} \dot{p}(x, t'; x_s) \int dt \dot{g}(x_r, t - t'; x, 0) \dot{p}(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs}.
\end{aligned} \tag{4.21}$$

As the Green's function is invariant by time translation, I have

$$\begin{aligned}
&\int dt \int dx_r p'(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs} \frac{\partial p(x_r, t; x_s)_{cal}}{\partial c(x)} = \\
&\int dt' \int dx_r \frac{-2}{c(x)^3} \dot{p}(x, t'; x_s) \int dt \dot{g}(x_r, 0; x, t' - t) p'(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs} = \\
&\int dt' \int dx_r \frac{-2}{c(x)^3} \dot{p}(x, t'; x_s) \dot{g}(x_r, 0; x, t') * p'(x_r + \Delta\sigma, t' + \Delta\tau; x_s)_{obs} = \\
&\int dt \int dx_r \frac{-2}{c(x)^3} \dot{p}(x, t; x_s) \dot{g}(x_r, -t; x, 0) * p'(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs}
\end{aligned} \tag{4.22}$$

and, similarly,

$$\begin{aligned}
&\int dt \int dx_r \dot{p}(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs} \frac{\partial p(x_r, t; x_s)_{cal}}{\partial c(x)} = \\
&\int dt \int dx_r \frac{-2}{c(x)^3} \dot{p}(x, t; x_s) \dot{g}(x_r, -t; x, 0) * \dot{p}(x_r + \Delta\sigma, t + \Delta\tau; x_s)_{obs}.
\end{aligned} \tag{4.23}$$

Sustituting these expressions into equations (4.16) and (4.17) I obtain

$$\begin{aligned}
\frac{\partial \Delta\sigma}{\partial c(x)} \Delta\sigma(x_s) = \frac{1}{c(x)^3} \Big[&\int dt \int dx_r \dot{p}(x, t; x_s) \dot{g}(x, -t; x_r, 0) * \delta\sigma \\
&- \int dt \int dx_r \dot{p}(x, t; x_s) \dot{g}(x, -t; x_r, 0) * \delta\tau \frac{d_2}{d_1} \frac{d_3}{d_4} \frac{\Delta\sigma}{\Delta\tau} \Big]
\end{aligned} \tag{4.24}$$

and

$$\frac{\partial \Delta\tau}{\partial c(x)}\Delta\tau(x_s) = \frac{1}{c(x)^3}\left[\int dt \int dx_r \dot{p}(x,t;x_s)\dot{g}(x,-t;x_r,0) * \delta\tau \right.$$
$$\left. - \int dt \int dx_r \dot{p}(x,t;x_s)\dot{g}(x,-t;x_r,0) * \delta\sigma \frac{d_1}{d_2}\frac{d_3}{d_5}\frac{\Delta\tau}{\Delta\sigma}\right], \tag{4.25}$$

with the pseudo-residuals $\delta\sigma$ and $\delta\tau$ defined as

$$\begin{aligned} \delta\sigma(x_r,t;x_s) &= -\frac{2}{d_1}p'(x_r+\Delta\sigma,t+\Delta\tau;x_s)_{obs}\Delta\sigma(x_s) \\ \delta\tau(x_r,t;x_s) &= -\frac{2}{d_2}\dot{p}(x_r+\Delta\sigma,t+\Delta\tau;x_s)_{obs}\Delta\tau(x_s). \end{aligned} \tag{4.26}$$

and

$$\begin{aligned} d_3 &= \int dt \int dx_r \dot{p}'(x_r+\Delta\sigma,t+\Delta\tau;x_s)_{obs}p(x_r,t;x_s)_{cal} \\ d_4 &= \int dt \int dx_r \ddot{p}(x_r+\Delta\sigma,t+\Delta\tau;x_s)_{obs}p(x_s,t;x_s)_{cal} \\ d_5 &= \int dt \int dx_r p''(x_r+\Delta\sigma,t+\Delta\tau;x_s)_{obs}p(x_r,t;x_s)_{cal}. \end{aligned} \tag{4.27}$$

Now γ_σ and γ_τ (see eq. (4.8)) can be written as

$$\gamma_\sigma = -\frac{1}{c(x)^3}\sum_s \left[\int dt \int dx_r \dot{p}(x,t;x_s)\dot{Q}(x,t;x_s) \right.$$
$$\left. - \int dt \int dx_r \dot{p}(x,t;x_s)\dot{R}(x,t;x_s)\frac{d_2}{d_1}\frac{d_3}{d_4}\frac{\Delta\sigma}{\Delta\tau}\right] \tag{4.28}$$

$$\gamma_\tau = -\frac{1}{c(x)^3}\sum_s \left[\int dt \int dx_r \dot{p}(x,t;x_s)\dot{R}(x,t;x_s) \right.$$
$$\left. - \int dt \int dx_r \dot{p}(x,t;x_s)\dot{Q}(x,t;x_s)\frac{d_1}{d_2}\frac{d_3}{d_5}\frac{\Delta\tau}{\Delta\sigma}\right], \tag{4.29}$$

with

$$\begin{aligned} \dot{Q}(x,t;x_s) &= \dot{g}(x,-t;x_r,0) * \delta\sigma(x_r,t;x_s) \\ \dot{R}(x,t;x_s) &= \dot{g}(x,-t;x_r,0) * \delta\tau(x_r,r;x_s), \end{aligned} \tag{4.30}$$

and

$$\begin{aligned} Q(x,t;x_s) &= -g(x,-t;x_r,0) * \delta\sigma(x_r,t;x_s) \\ R(x,t;x_r) &= -g(x,-t;x_r,0) * \delta\tau(x_r,t;x_s). \end{aligned} \tag{4.31}$$

On the other hand, the result

$$\frac{d_2}{d_1} = \left(\frac{d_1}{d_2}\right)^{-1} = \frac{\int dt \int dx_r \ddot{p}_{obs} p_{cal}}{\int dt \int dx_r p''_{obs} p_{cal}} , \tag{4.32}$$

permits equations (4.28) and (4.29) to be rewritten as

$$\begin{aligned} \gamma_\sigma = -\frac{1}{c(x)^3} \sum_s \Big[&\int dt \int dx_r \dot{p}(x,t;x_s) \dot{Q}(x,t;x_s) \\ &- \int dt \int dx_r \dot{p}(x,t;x_s) \dot{R}(x,t;x_s) \frac{d_3}{d_5} \frac{\Delta\sigma}{\Delta\tau} \Big] \end{aligned} \tag{4.33}$$

$$\begin{aligned} \gamma_\tau = -\frac{1}{c(x)^3} \sum_s \Big[&\int dt \int dx_r \dot{p}(x,t;x_s) \dot{R}(x,t;x_s) \\ &- \int dt \int dx_r \dot{p}(x,t;x_s) \dot{Q}(x,t;x_s) \frac{d_3}{d_4} \frac{\Delta\tau}{\Delta\sigma} \Big] . \end{aligned} \tag{4.34}$$

Finally, the steepest descent direction of the misfit function S is given by

$$\begin{aligned} \gamma = \gamma_\sigma + \beta\gamma_\tau = \\ -\frac{1}{c(x)^3} \sum_s \Big[&\int dt \int dx_r \dot{p}(x,t;x_s) \dot{Q}(x,t;x_s) \left[1 - \beta \frac{d_3}{d_4} \frac{\Delta\tau}{\Delta\sigma} \right] \\ &+ \int dt \int dx_r \dot{p}(x,t;x_s) \dot{R}(x,t;x_s) \left[\beta - \frac{d_3}{d_5} \frac{\Delta\sigma}{\Delta\tau} \right] \Big] , \end{aligned} \tag{4.35}$$

where $Q(x,t;x_s)$ and $R(x,t;x_s)$ represent the back-propagated pressure fields calculated with the pseudo-residuals $\delta\sigma$ and $\delta\tau$ acting as sources at the receiver locations x_r. I remark on the similarity between the two terms of this expression and the analogous equation in the standard wave-equation traveltime inversion method, equation (2.47).

4.2 Summary

This analysis describes the procedure for finding a new velocity model using a bi-dimensional cross-correlation function in the time and offset domain as a connective function. The proposed method takes into account the moveout between traces, whereas the standard traveltime inversion treats adjacent traces as independent. I define the new bi-dimensional misfit function as the sum of the squared offset and traveltime delays. This function has a minimum when the calculated and the observed shot gathers best match each other. I use the steepest descent method to update the velocity values in an iterative way.

The steepest descent direction of the misfit function is calculated using the connective function, the Fréchet derivative, and properties of the Green's function. The velocity field is then updated by correlating at zero lag the time derivative of the current wavefield with the time derivative of the pressure field calculated by backward propagation of two pseudo-residuals acting as sources at the receiver locations.

Chapter 5

Wave-equation traveltime-offset inversion: numerical aspects

In general the structure of the wave-equation inversion codes is similar, however there are slight differences in one method with respect to the others. For instance, all of them correlate a forward wavefield with a back-propagated wavefield, but the sources for the reverse time propagation differ from each other. For the calculation of the pseudo-residuals the full waveform inversion does not require the cross-correlation code, the standard traveltime inversion uses the one-dimensional cross-correlation function, and the new traveltime-offset inversion needs a bi-dimensional cross-correlation code.

On the other hand, the values of the preconditioning and the mute functions are not directly related to the inversion method, but to the geometry configuration of the data to be inverted. In this chapter, I give a general overview of the components that build up the wave-equation methods. Each section defines a particular step.

5.1 The parallel computer code

As I mentioned before, a standard workstation computer code for full waveform inversion was provided by the University of Utah Tomography Consortium. I modified this original program to execute the standard traveltime and the traveltime-offset inversions. I generated the parallel computer codes, and I implemented them in the Cray T3D computer at the Edinburgh Parallel Computing Centre (EPCC). The parallel versions use the Message-Passing Interface (MPI) (Message Passing Interface Forum, 1994).

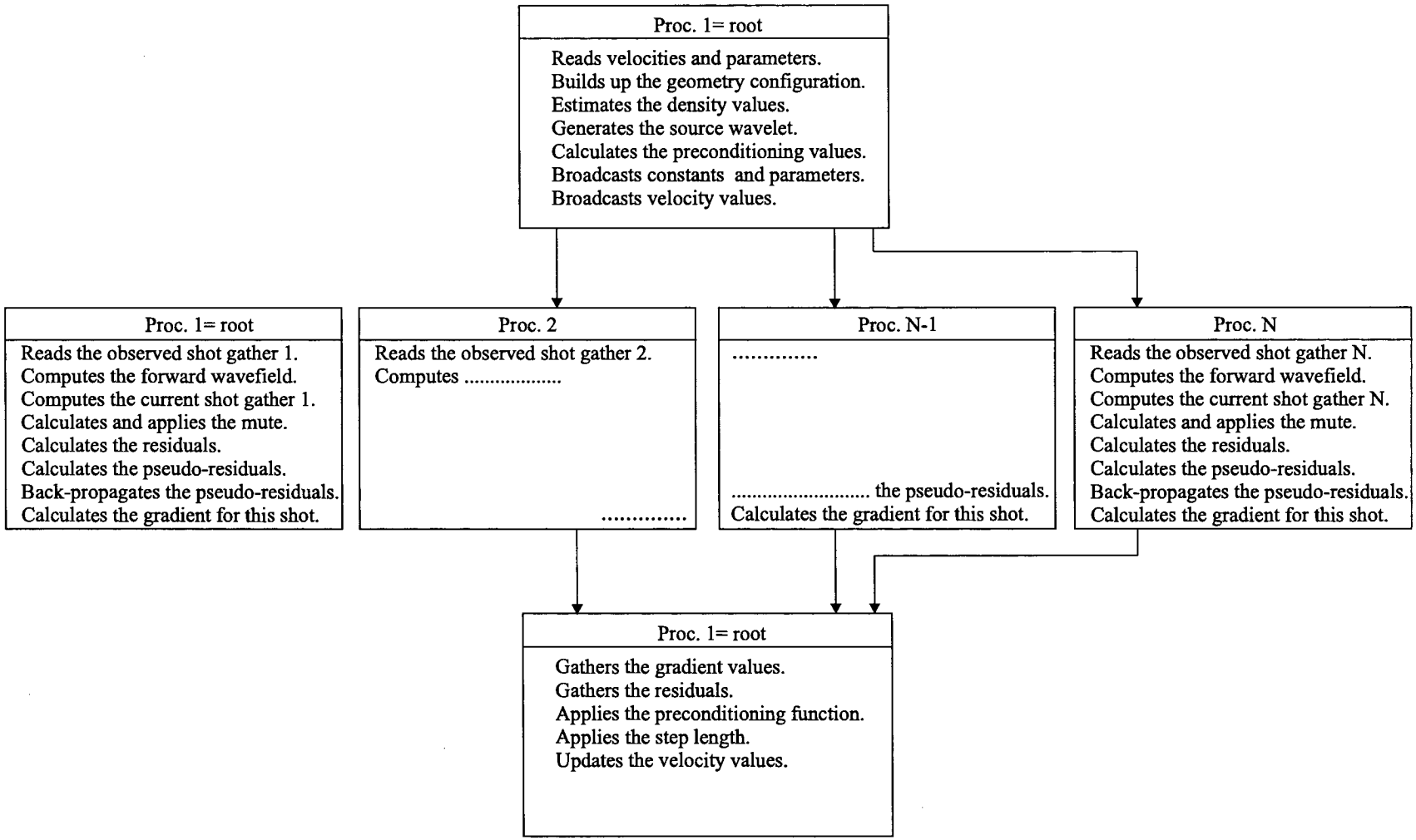


Figure 5.1: Scheme describing the parallel code.

Figure 5.1 shows the scheme for the parallel code. The processor "root" commands the execution of the inversion, reads the input parameters and the initial velocity field, computes the constant values needed for the inversion, and distributes the required information to all processors. Each individual processor forward and backward propagates the corresponding wavefields, and computes the contribution of a particular shot gather to the gradient and to the residuals. The processor "root" gathers and adds up the calculated gradients. It also sums the residuals, and multiplies the total gradient by the step length value and the preconditioning function. Finally, it updates the velocity values.

The execution CPU time varies with the inversion method. To process a dataset with 15 shots, 62 traces, 1000 samples in time, and a grid of 129x150 points, the full waveform inversion takes 340 s of CPU time on the Cray T3D for each iteration, and the standard traveltime inversion needs approximately 400 s. Meanwhile, the traveltime-offset inversion requires 770 s. The computer code for the last inversion has not been optimised yet. It is not optimum for two reasons: first, the cross-correlograms are calculated using a discrete form of equation (4.2) instead of using the Fourier Transform, and secondly each individual processor is in charge of back-propagating two different pseudo-residuals ($\delta\sigma$ and $\delta\tau$). I expect a lower execution time after optimising the computer code.

5.2 The forward and backward propagations

A staggered grid finite-difference scheme (Luo and Schuster, 1991; Savic, 1995) is used for the forward modelling and for the reverse-time propagation. In order to implement this scheme, Luo and Schuster use the Newton's second law of motion and the Hooke's law to rewrite the wave equation

$$\frac{1}{c(x)^2}\frac{\partial^2 p(x_r,t;x_s)}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla p(x_r,t;x_s)\right] = s(t;x_s), \qquad (5.1)$$

as

$$\begin{aligned} \frac{\partial p(x_r,t;x_s)}{\partial t} &= c(x)^2\rho(x)\nabla\cdot[w(x_r,t;x_s)] + c(x)^2\int dt\, s(t;x_s), \\ \frac{\partial w(x_r,t;x_s)}{\partial t} &= \frac{1}{\rho(x)}\nabla p(x_r,t;x_s), \end{aligned} \tag{5.2}$$

for the forward propagation. Similarly, for the reverse-time propagation the wave equation

$$\frac{1}{c(x)^2}\frac{\partial^2\dot{q}(x_r,t;x_s)}{\partial t^2} - \rho(x)\nabla\cdot\left[\frac{1}{\rho(x)}\nabla\dot{q}(x_r,t;x_s)\right] = \dot{\delta\tau}(x_r,t;x_s), \tag{5.3}$$

can be written as

$$\begin{aligned} \frac{\partial\dot{q}(x_r,t;x_s)}{\partial t} &= c(x)^2\rho(x)\nabla\cdot[\dot{u}(x_r,t;x_s)] + c(x)^2\delta\tau(x_r,t;x_s), \\ \frac{\partial\dot{u}(x_r,t;x_s)}{\partial t} &= \frac{1}{\rho(x)}\nabla\dot{q}(x_r,t;x_s), \end{aligned} \tag{5.4}$$

Here w and u are the particle velocity vectors.

The initial conditions are given by

$$p(x_r,0;x_s) = 0; \qquad w(x_r,0;x_s) = 0, \tag{5.5}$$

$$\dot{q}(x_r,T;x_s) = 0; \qquad \dot{u}(x_r,T+1/2\Delta t;x_s) = 0, \tag{5.6}$$

where T and Δt are the total recording length and the time-step, respectively.

I use these equations to implement the standard traveltime inversion, applying the same procedure for the back-propagation of the two pseudo-residuals in the traveltime-offset inversion. The derivatives are replaced by operators of second order in time and fourth order in space. The stability condition for this finite-difference operator is

$$c_{max}\Delta t \leq 0.6\Delta x. \tag{5.7}$$

with Δx the grid size. Also 5 grid points are considered for the shortest wavelength:

$$5\Delta x \leq c_{min}/\nu_{max}. \tag{5.8}$$

The computer code considers free surface and absorbing boundary conditions. In my numerical examples, I use four absorbing boundaries (Savic, 1995; Zhou, 1995; Zhou et al., 1995) to minimise unwanted reflections at the edges of the model.

5.3 The cross-correlation programs

The conventional traveltime inversion has as the connective function the one-dimensional cross-correlation function described in the time domain by equation (2.18). The computing program uses the cross-correlation theorem, equation (2.49), to accelerate the calculation of the traveltime delays. The Fast Fourier Transform subroutine (rfft) is used, and the traces are correlated independently. There are as many shifts in time as traces in the shot gather.

The moveout between traces is maintained during the traveltime-offset inversion. The bi-dimensional cross-correlation function, equation (4.2), moves the complete observed shot gather over the calculated one, giving only one shift in time as well as one shift in offset for each shot gather. Even though the bi-dimensional Fourier transform can be used, the current computer code executes this process using the discrete form (Gonzalez and Wintz, 1977; Marion, 1991) given by

$$f(\sigma,\tau;x_s) = \frac{1}{h}\sum_{x=0}^{nx-1}\sum_{t=0}^{nt-1} p(x+\sigma,t+\tau;x_s)_{obs}p(x,t;x_s)_{cal}, \tag{5.9}$$

with

$$-(nx-1) \leq \sigma \leq (nx-1), \quad -(nt-1) \leq \tau \leq (nt-1). \tag{5.10}$$

Here nx and nt are the number of traces in the shot gather and the number of samples in time, respectively. The weighting factor h is computed by counting the number of samples in offset and time that contribute to a particular calculation. In practice the shifts and delays must be small compared with the offset and time windows.

The user can define a cross-correlation window that restricts the program to compute values within the offset and time delays enclosed in that window, i.e., the observed shot gather is shifted over the calculated shot gather up to $\pm\Delta\sigma_{max}$ in offset and up to $\pm\Delta\tau_{max}$ in time. This helps to avoid secondary maxima by controlling the relative movement between the shot gathers, and also helps to reduce the CPU time of the calculations. Figure 5.2 shows the cross-correlograms where two different windows are used for the calculations.

5.4 The mute functions

Rowbotham and Goulty (1993) suggest three different methods for removing the first arrivals: muting, estimation and subtraction, and f-k filtering. I consider only the first method. The function implemented by the University of Utah Tomography Consortium to carry out the mute of the direct arrivals in their faulted synthetic model, is given by

$$m_c(it) = \begin{cases} 0 & \text{if } it < it_{max} - 40, \\ \frac{1}{2}\left[1 + \tanh[\frac{(it_{max} - it)}{4}]\right] & \text{if } it_{max} - 40 \leq t \leq it_{max} + 80, \\ 1 & \text{if } it > it_{max} + 80, \end{cases} \tag{5.11}$$

with it_{max} the time sample where the trace has the maximum amplitude, i.e., the peak of the direct arrival.

I modified this function in order to consider datasets with surface reflection geometries,

$$m_s(it) = \frac{1}{2}\left[1 + \tanh[\frac{(it - mutt)}{amut} - cmut]\right], \tag{5.12}$$

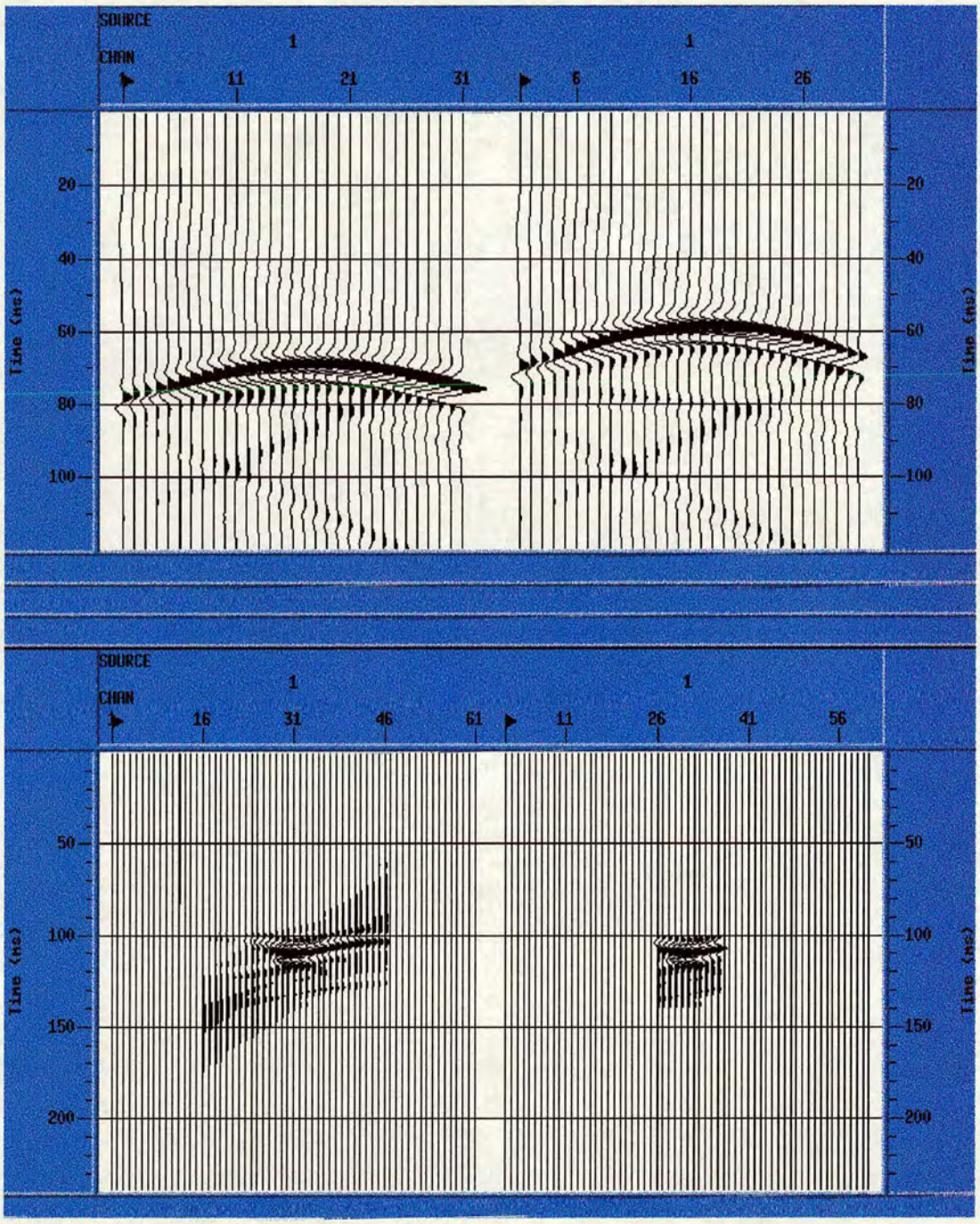


Figure 5.2: Top: shot gathers. Bottom: cross-correlograms using two different windows ($\Delta\sigma_{max} = 15, \Delta\tau_{max} = 300$, and $\Delta\sigma_{max} = 5, \Delta\tau_{max} = 100$, respectively).

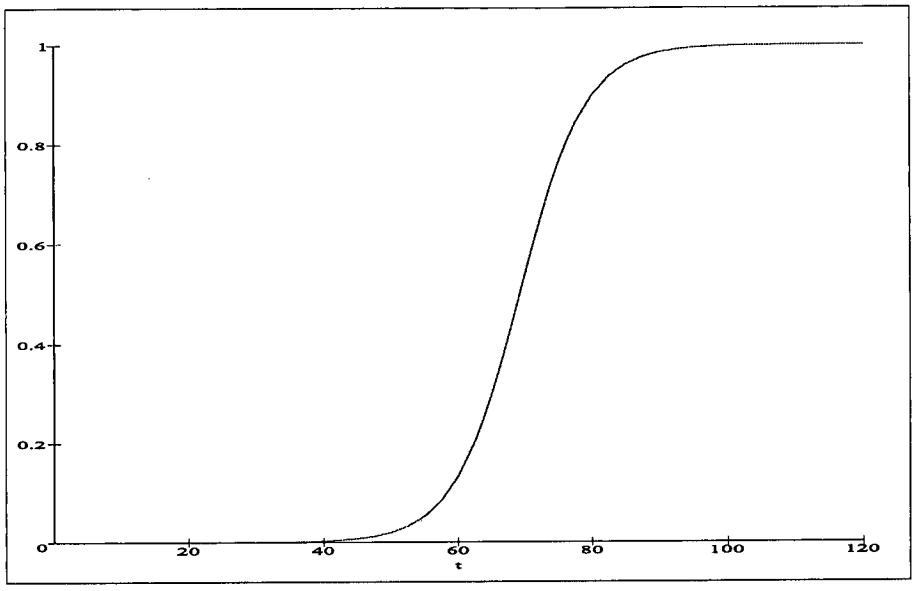


Figure 5.3: Hyperbolic mute function.

with

$$mutt = it_{max} + nts, \tag{5.13}$$

$$cmut = \tanh^{-1}(\frac{2}{nfa} - 1), \tag{5.14}$$

$$amut = \frac{nts}{(\tanh^{-1}(2 * 0.98 - 1) - cmut)}, \tag{5.15}$$

$$nts = (\nu_{peak}\Delta t)^{-1}. \tag{5.16}$$

The variable *amut* defines the slope of the hyperbolic function by assigning 98% of the original amplitude to the sample in the time position $it_{max} + 2nts$. *cmut* uses the attenuation factor nfa to modify the amplitude of the trace at the time $it = mutt$. For instance, if $nfa = 100$ the amplitude of the muted trace at $it = it_{max} + nts$ is 0.01 times the original amplitude. Figure 5.3 displays the shape of this mute function.

5.5 The preconditioning function

As mentioned by Gauthier et al. (1986), Mora (1988), and Luo and Schuster (1991), a preconditioning function is required to compensate for the irregular distribution of the energy in the gradient, and to accelerate the convergence. I use a preconditioning function that corrects for the geometrical expansion of the energy (Luo and Schuster, 1991), thus speeding up the convergence. It is given by

$$\Phi(x) = \|x - x_s\|^{1/2} \|x - x_r\|^{1/2} \tag{5.17}$$

As the position of the sources and the receivers depends upon the geometry configuration of the dataset, this preconditioning function varies for the same point x. Less energy reaches the top and the bottom of the model when using crosswell data, while the energy has a maximum value at the top of the model when sources and receivers are positioned on the surface. Figure 5.4 shows the maps of the function for the crosshole and the surface reflection configurations, respectively.

5.6 The smoothing functions

Two different smoothing functions are available in the traveltime-offset computer code. The first one corresponds to a bi-dimensional Gaussian filter in the (k_x, k_z) domain. Applying the convolutional theorem (Bracewell, 1986; Brook and Wynne, 1988), the current velocity model is Fourier transformed to the mentioned domain, multiplied by the Gaussian function in that domain, and inverse Fourier transformed back to the (x, z) domain.

The second smoothing function (α-trimmed means) was proposed by Gersztenkorn and Scales in 1988. It is defined by

$$Y_{\alpha_s} = \frac{1}{N - 2j} \sum_{i=j+1}^{N-j} Y_i, \tag{5.18}$$

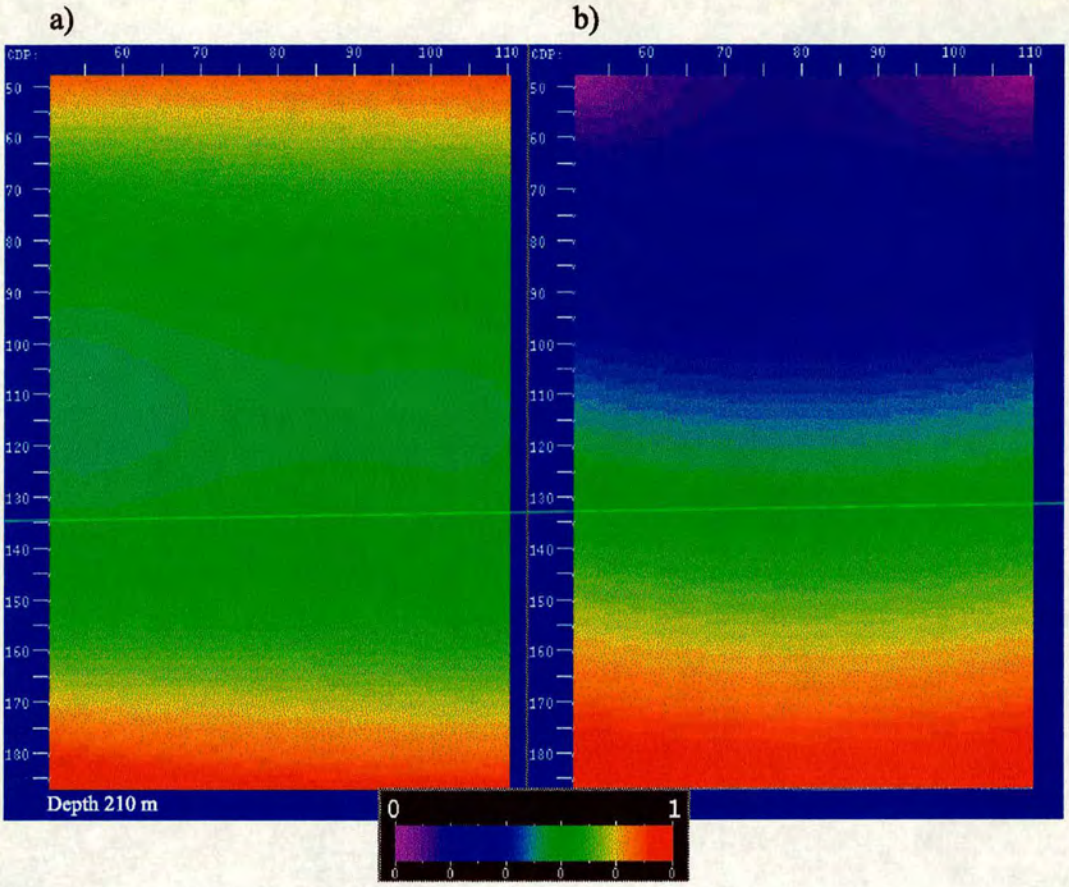


Figure 5.4: Preconditioning function. a) For crosshole datasets. b) For surface reflection geometries.

where

$$j = \begin{cases} \alpha_s N & N \text{ odd}, \\ \alpha_s (N-1) & N \text{ even}, \end{cases} \tag{5.19}$$

with $0 \leq \alpha_s \leq 0.5$.

In this equation Y_i is an array of velocity values in N nodes (9 in this computer code) arranged in ascending or descending order, and α_s is the "adjuster". When α_s varies from $\alpha_s = 0$ and $\alpha_s = 0.5$, Y_{α_s} oscillates between the mean and the median values, respectively. The median filter helps to keep the borders and the lateral structures, while the mean filter smooths the data in the RMS sense.

5.7 Summary

I present a general overview of different steps that are involved in the wave-equation inversion methods. The forward and reverse time propagation is carried out using a finite-difference scheme of second order in time and fourth order in space. The cross-correlation theorem is used to determine the traveltime delays in the standard traveltime inversion. In the new inversion method, the traveltime and the offset delays are computed using a discrete form of the bi-dimensional cross-correlation function.

A hyperbolic tangent function is used to mute out the direct arrivals. The preconditioning function speeds up the convergence of the inversions by correcting the irregular distribution of the seismic energy. Two different smoothing functions have been set up to remove the unwanted high frequency features present in the current velocity model. Finally, the parallel codes of the inversions, which use the MPI technique, are implemented on a Cray T3D computer at the Edinburgh Parallel Computing Centre.

Chapter 6

Traveltime-offset inversion: Examples and discussion

The inversion of seismic datasets is not a straightforward problem. The procedure involves many different parameters whose values can considerably affect the results. Some authors have described, for instance, the behaviour of the full waveform inversion when the ultra long wavelength features of the velocity model are not present in the initial model. Others show the influence of density variations as well as the addition of noise to the datasets. They report that inverting real datasets introduces further complications as it requires the estimation of the source wavelet.

It is not only physical parameters which control the inversion; the numerical simulation also requires extra parameters to achieve the desired convergence. Among them, the step length and the preconditioning factors play an important role. I deal with all of them when inverting different synthetic datasets using the proposed traveltime-offset inversion method. Even though the offset inversion allows a more stable convergence by taking into account the moveout, the difficulties of the inversion are increased by the fact that I use, in most of the examples, traveltime information from reflection events.

6.1 Inversion of a crosshole dataset

I consider again the crosshole dataset used by Luo and Schuster in 1991, and Zhou et al. in 1995 (see Figure 6.1). The geometry and parameters for the inversion are fully described in Chapter 2 with the exception that 18 shots instead of 15 are placed in the left hand side of the model. The complete seismograms

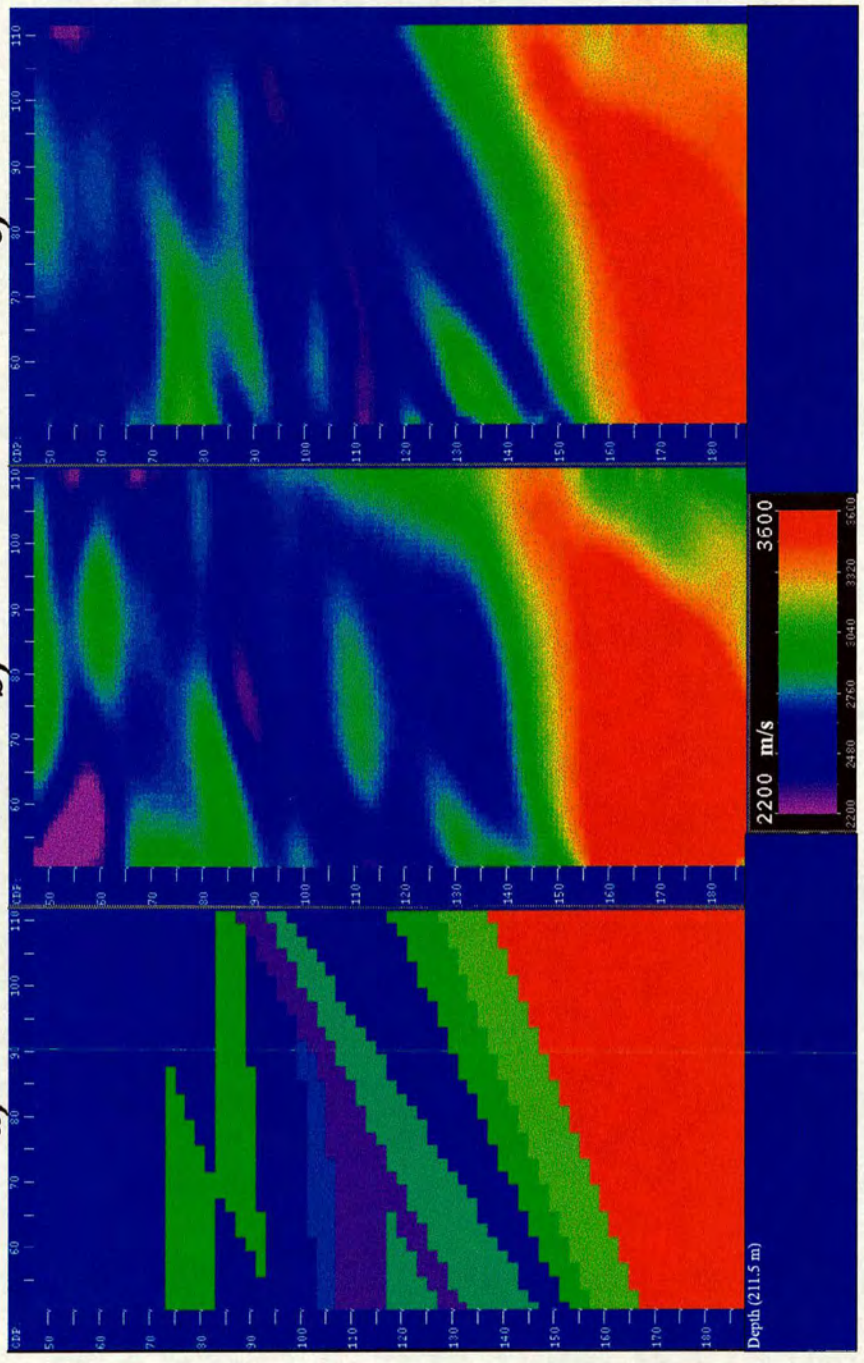


Figure 6.1: Inversion of the fault model used in Luo and Schuster (1991). a) Fault model (crosshole dataset). b) Traveltime-offset inversion after 3 iterations. c) Full waveform inversion after 7 iterations, starting with the velocity model b).

are used for the two-step inversion, i.e. the mute is not applied. The β value that keeps the dimensions consistent is set up to 1.0 m^2/s^2. The traveltime and offset delays are calculated using the bi-dimensional cross-correlation function.

In the first step, three iterations are executed using the traveltime-offset inversion method, and the low frequency content of the actual velocity model is recovered (Figure 6.1b). This output becomes the starting model for the full waveform inversion, the second step, that incorporates the high frequency features in the result. After 7 iterations the final reconstruction, shown in Figure 6.1c, is comparable to the result from the standard traveltime + full waveform inversion displayed in Figure 2.5. Only 3 iterations are performed to get a low frequency image using the new method, while 10 iterations are needed for the standard inversion. Figures 6.2 and 6.3 show the offset and time delays for each shot. Although the residuals are not null (see white squares in the figures) the convergence is stable.

Due to the fact that the hybrid method, traveltime-offset and full waveform inversion, succeeds in the reconstruction of a moderately complex fault model, I do not consider additional examples with a crosshole configuration. Instead of that I dedicate more effort to the traveltime and offset inversion of datasets with surface reflection geometry.

6.2 The point-diffractor

First a simple model is considered, which consists of a point diffractor embedded in a constant velocity medium of 2000 m/s. A set of 15 shots and 31 receivers is evenly spread on the top of the model to simulate the surface reflection geometry. Even though I display only 77 nodes in depth in the figures, a grid of 62x141 points (1.5 m apart) is used to digitise the velocity model. The source function is a Ricker wavelet which has 120 Hz as a peak frequency. The first arrivals are muted out before the cross-correlation.

For this example the minimum velocity permitted in the inversion is 2000 m/s,

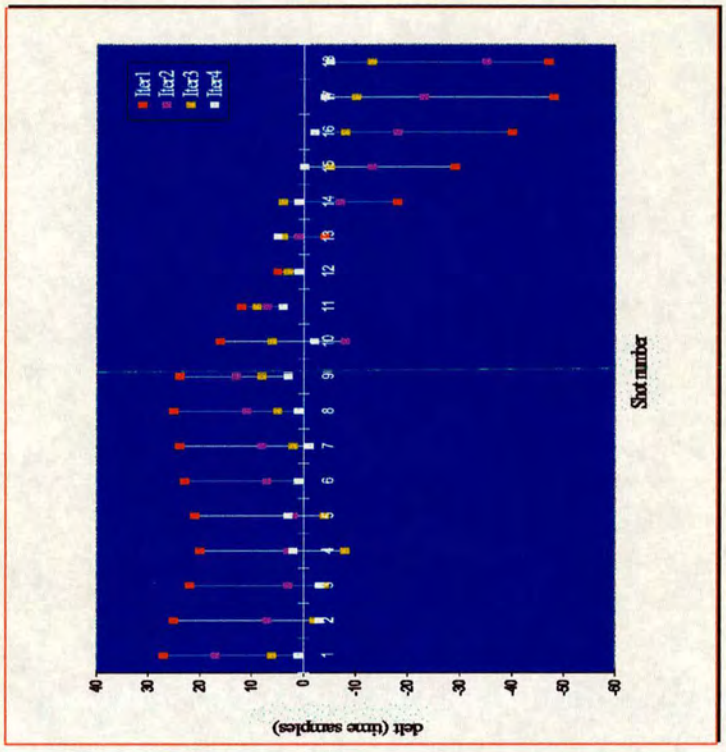


Figure 6.3: Traveltime delays for the crosshole dataset.

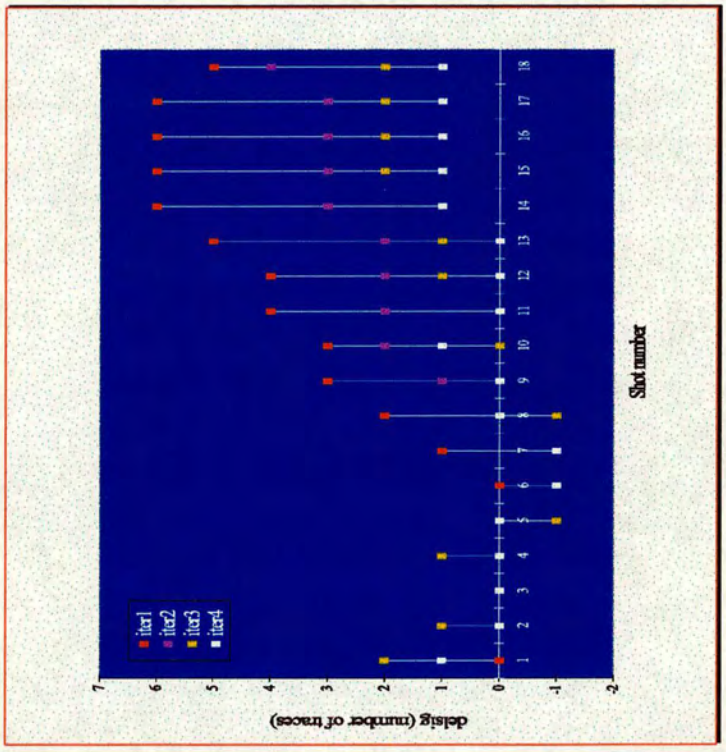


Figure 6.2: Offset delays for the crosshole dataset.

and for the starting model I changed the position of the original diffractor. The results are shown in Figure 6.4 for 1, 10 and 20 iterations. The point-diffractor is located at the right depth, and the convergence is stable. The delays in time and offset are zero for the velocity model emerging after the 20th iteration. The standard traveltime inversion gets a similar result but with prominent artifacts and unstable convergence, see Figure 2.16.

A second model with a point-diffractor is introduced to evaluate the lateral resolution of the traveltime-offset inversion in relation to depth. The new point-diffractor is located 12 m above the previous one, and a Ricker wavelet with 180 Hz as a peak frequency is used. Now, I place the diffractor in the starting velocity model beside the original point-diffractor, 12 m apart to the right. Figure 6.5 depicts the results for both original models. It is difficult to conclude that the shallower diffractor (Figure 6.5b) is less extended laterally than the deeper diffractor just by looking at them, but the artifacts above and below them seem to indicate that the lateral resolution decreases with depth. In contrast, there is a remarkable difference between the lateral resolution of the traveltime-offset inversion and the full waveform inversion. Figure 6.6 displays the results using a shallower point-diffractor as a starting model in both inversions. The vertical edges are well defined after full waveform inversion.

Figures 6.7a and 6.7b show the inversion of the point-diffractor with an array of 62 and 31 geophones, respectively. In the first array there is one receiver per node, while in the second one there is one receiver in every other node. There are not important differences between these particular cases. However, a change in the peak frequency from 120 Hz to 180 Hz, figures 6.7b and 6.7c, improves the resolution of the point-diffractor in depth. The resolution is approximately $\lambda_\nu/4$, with λ_ν the wavelength of the peak frequency.

6.3 Two point-diffractors

In surface reflection geometries, usually there is more than one reflected or diffracted arrival and the cross-correlation function is dominated by the event with major

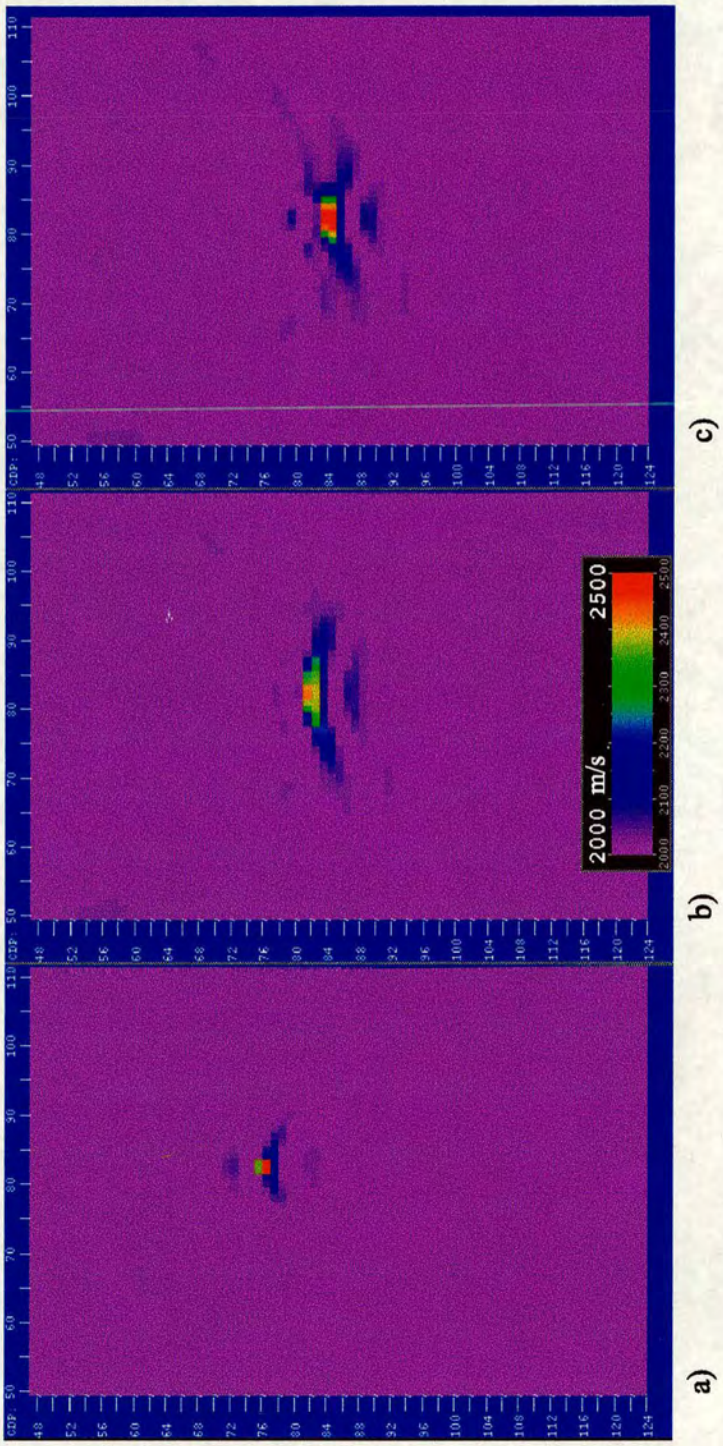


Figure 6.4: Traveltime-offset inversion of the point-diffractor model. a) After 1 iteration. b) After 10 iterations. c) After 20 iterations.

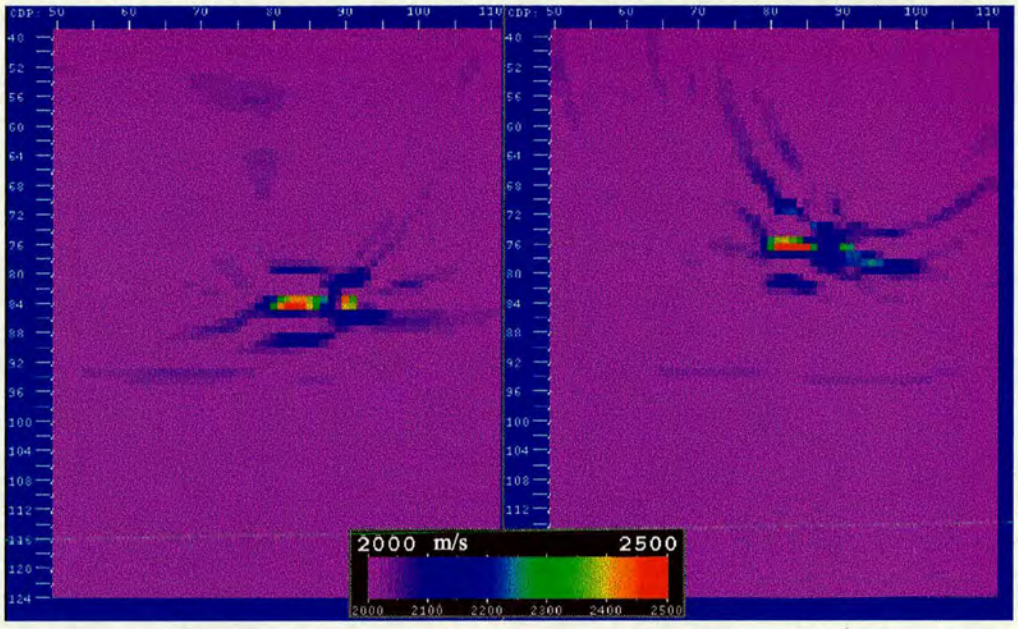


Figure 6.5: Evaluation of the lateral resolution. Traveltime-offset inversions with the original point-diffractor located at: a) 54 m depth, b) 42 m depth.

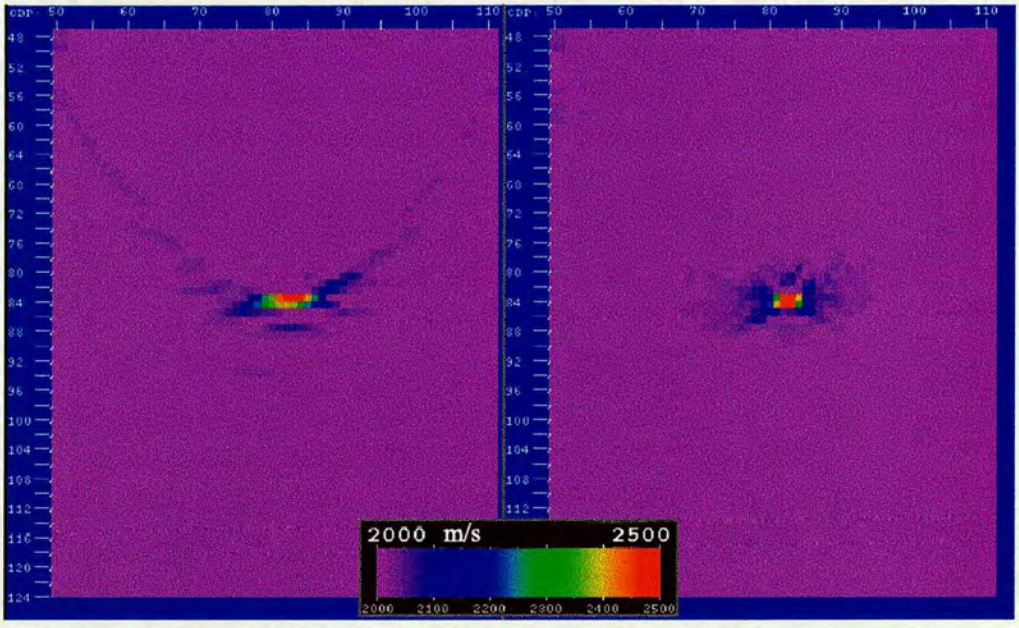


Figure 6.6: Lateral resolution of the wave-equation inversions. a) Traveltime-offset inversion. b) Full waveform inversion.

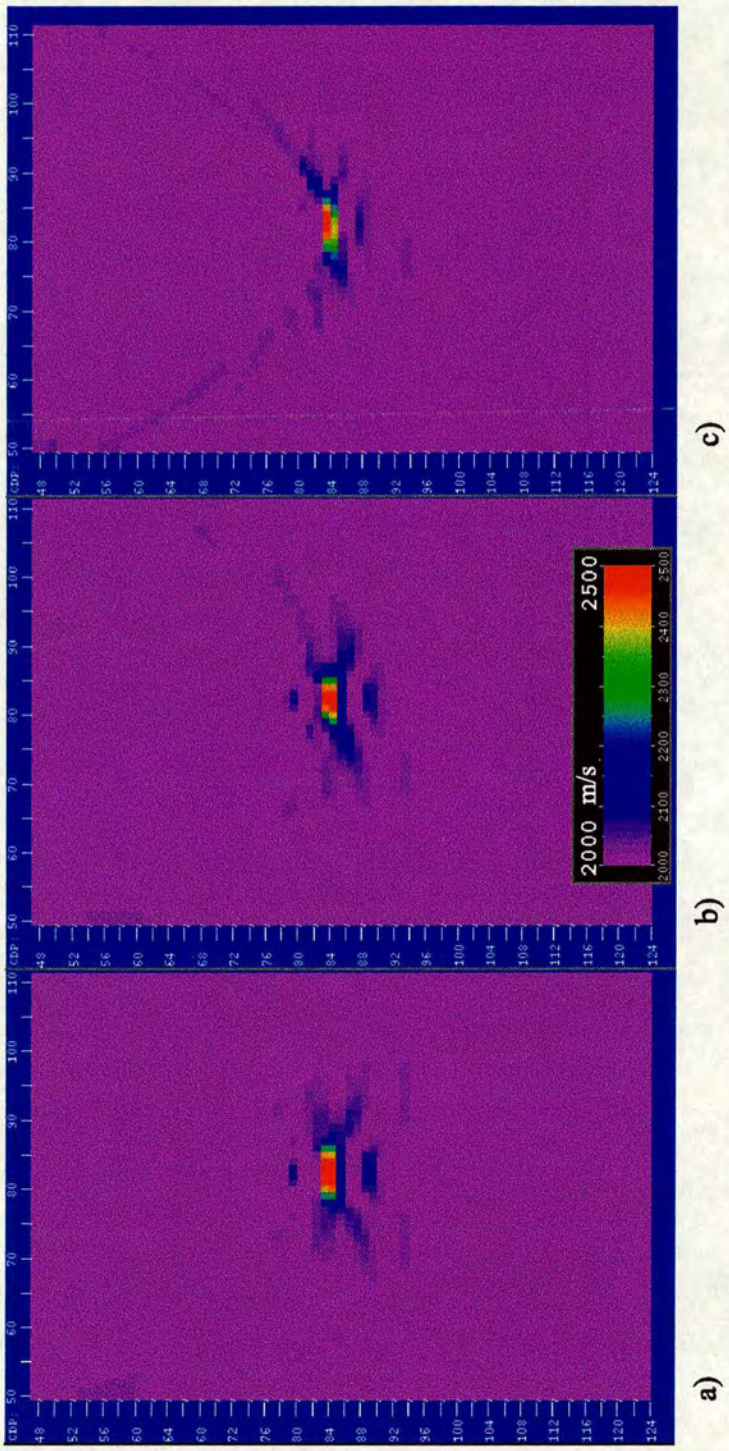


Figure 6.7: Traveltime-offset inversion of the point-diffractor model using different arrays, and different peak frequencies. a) 62 receivers and 120 Hz. b) 31 receivers and 120 Hz. c) 31 receivers and 180 Hz.

amplitudes. Inversions having the cross-correlation function as the connective function give velocity and depth values that depend greatly on the traveltime of this particular event. Later I will show how the inversion handles the additional information contained in the common shot gathers in complex models.

In the case of two point-diffractors sharing the same background velocity, the standard traveltime and the traveltime-offset inversions manage to locate the second point-diffractor in the right position although the first diffractor is dominant (see Figure 6.8). The back-propagated wavefield carries the information about the second diffractor and, as the background velocity is correct, it is placed in the right depth.

In the standard traveltime inversion the result is unstable, there are unexpected spikes in the time residuals and artifacts are present all around the velocity model. The traveltime-offset inversion succeeds after 24 iterations. The offset and time residuals decrease to zero as shown in Figure 6.9. The calculated shot gathers (Figure 6.10) show a better match to the observed shot gathers than the resulting shot gathers from the standard traveltime inversion, especially the shots by the sides of the model. The number of iterations in the traveltime-offset inversion can be reduced by increasing the step length to a bigger value in the first iteration.

In Chapter 2, Figure 2.18, I compare the results of the standard traveltime method when inverting the two point-diffractor velocity model with two different model depths. Figure 6.11 depicts the outputs from the traveltime-offset inversion. Figure 6.11a corresponds to a maximum depth of 210 m (141 nodes), whereas Figure 6.11b has a maximum depth of 90.5 m (62 nodes). It can be seen here that the lower absorbing boundary does not have an important influence in the results, and in both cases the convergence is stable. These results are a considerable improvement over the ones obtained using the standard traveltime inversion.

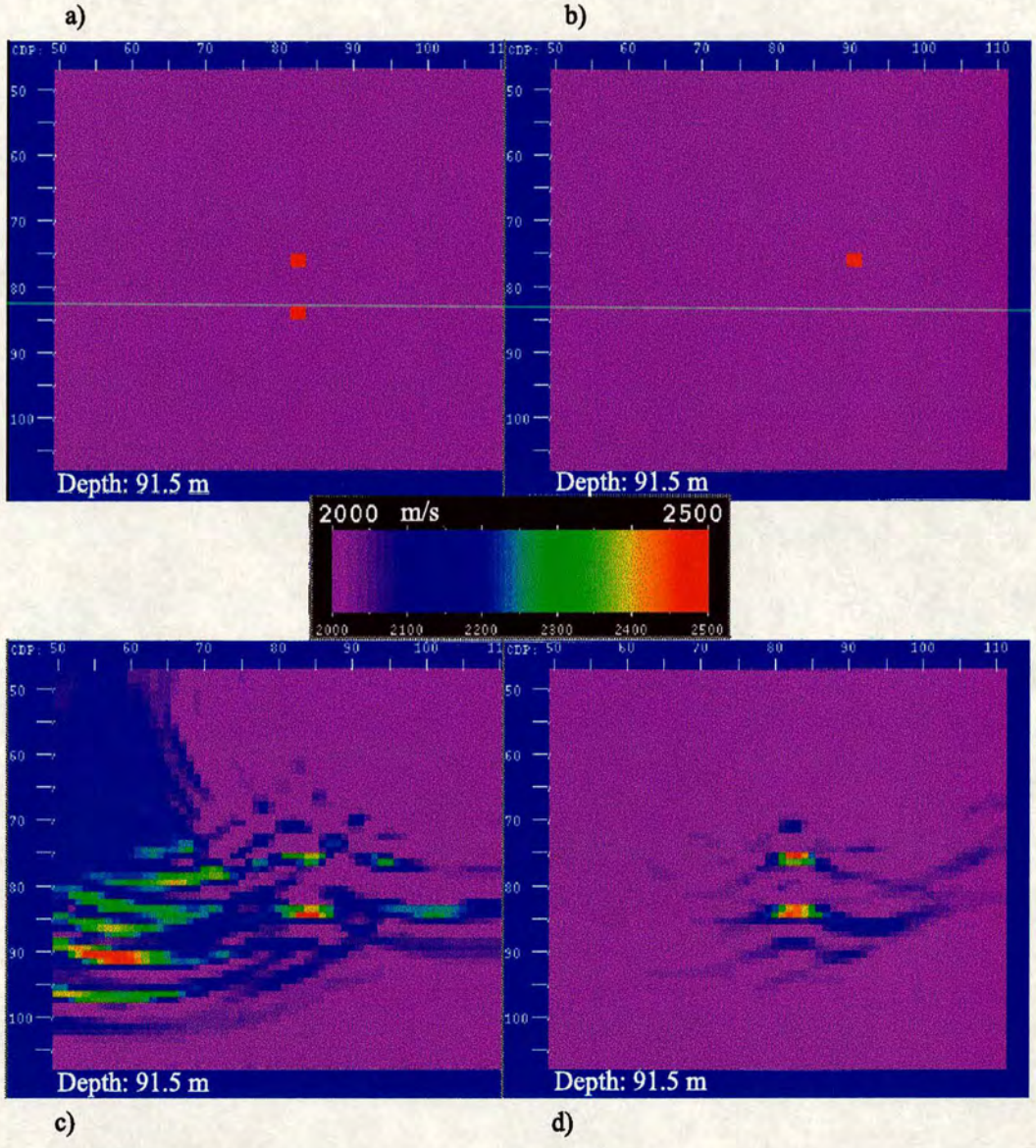


Figure 6.8: Inversion of the two point-diffractor model. a) Original model. b) Initial guess. c) Standard traveltime inversion. d) Traveltime-offset inversion.

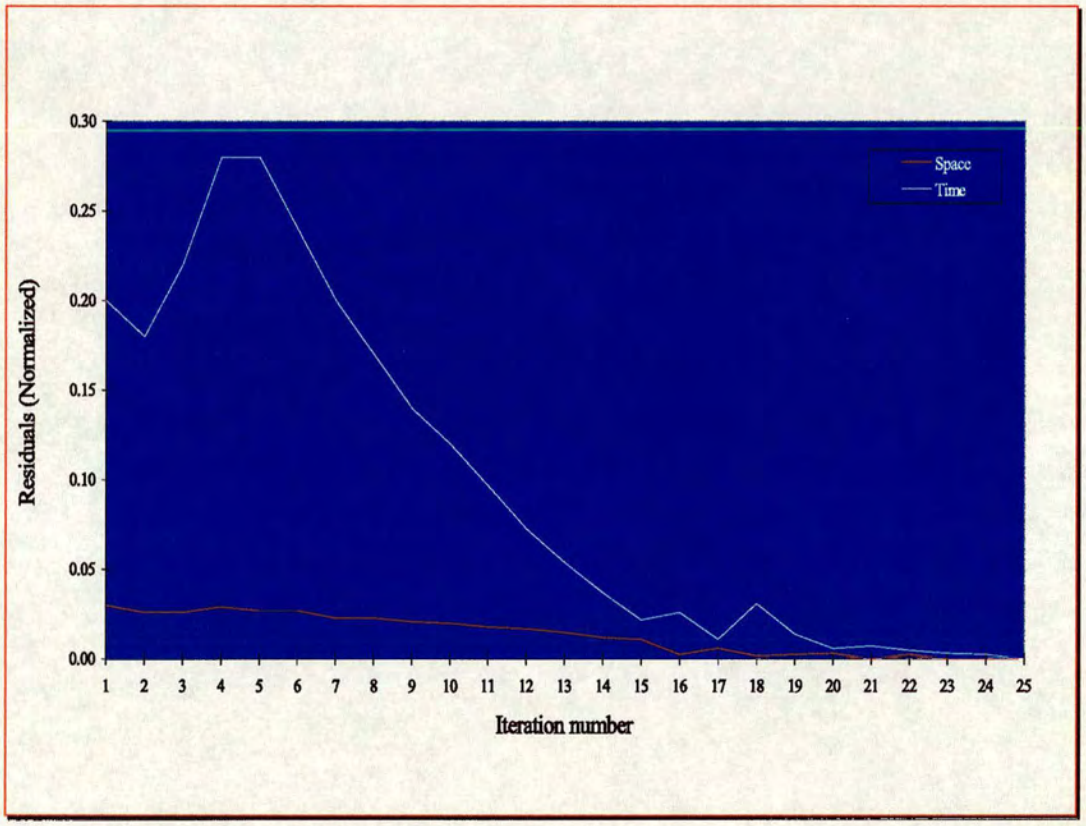


Figure 6.9: Residuals from the traveltime-offset inversion.

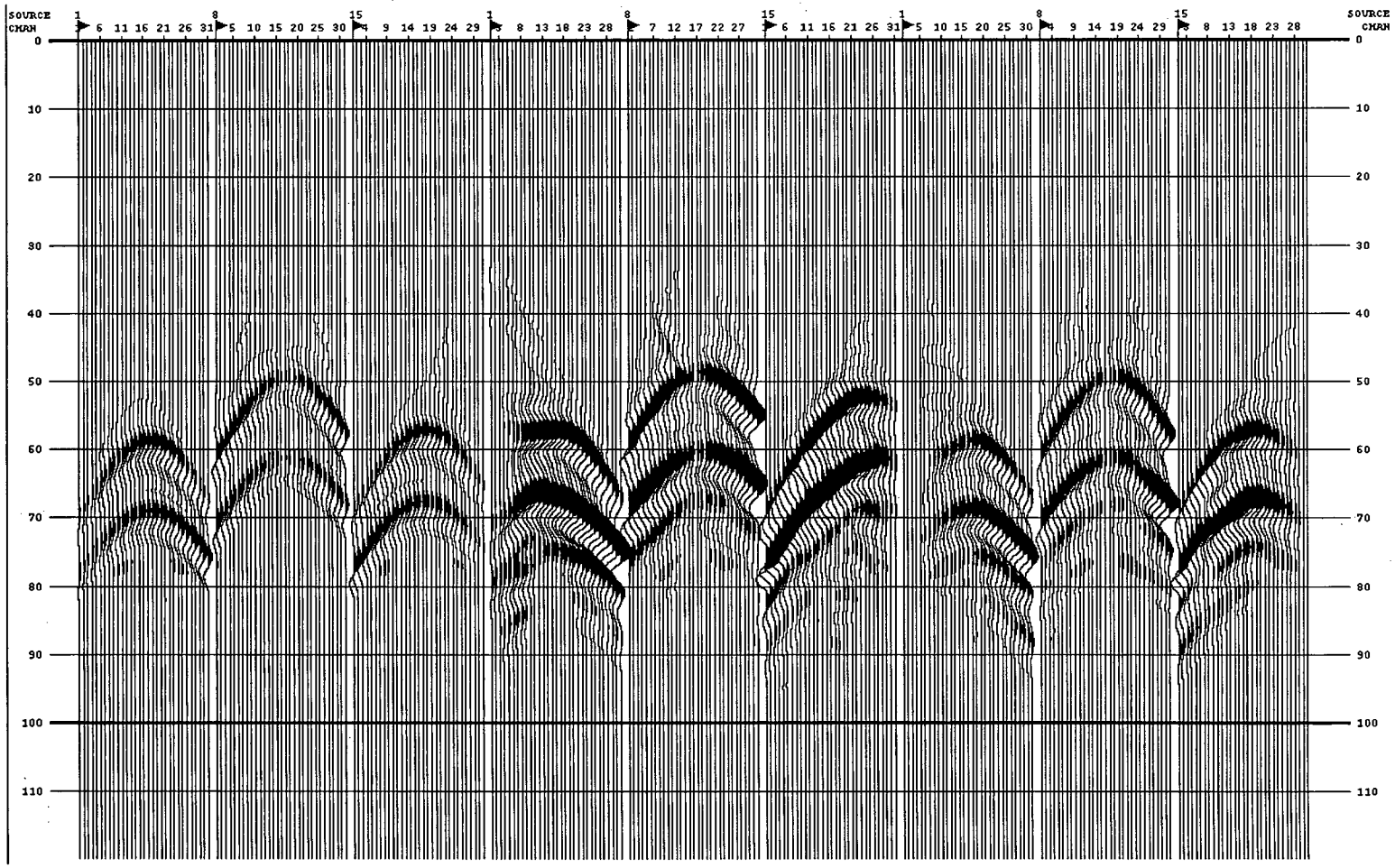


Figure 6.10: Shot gathers from the inversion of the two point-diffractor model. Left: original shot gathers. Centre: shot gathers from the standard traveltime inversion. Right: shot gathers from the traveltime-offset inversion.

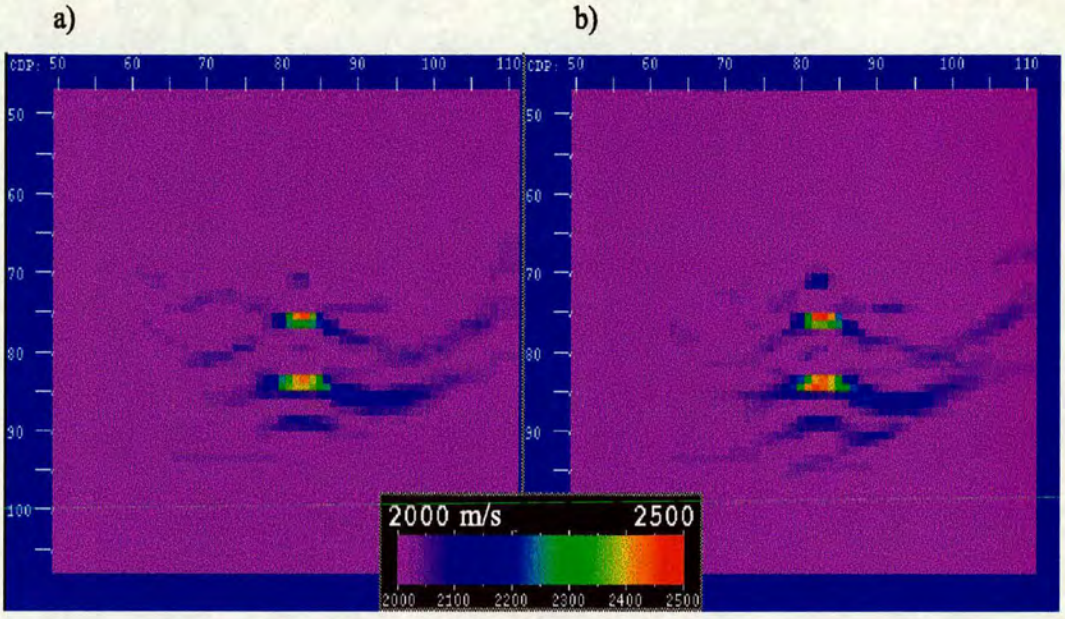


Figure 6.11: Traveltime-offset inversion of the same model, but placing the lower boundary at different depth: a) 210 m, b) 90.5 m.

6.4 Two point-diffractors shifted laterally

Rather than keeping the point-diffractors parallel in depth, I now displace the second one 12 m to the right. The parameters for the inversion are the same as in the preceding examples. The starting velocity model is a single point-diffractor located 12 m apart from the first diffractor, and 12 m above the second one in a medium with 2000 m/s as a background velocity.

Determining the value of the step length is not an easy matter. There are analytical estimations of it, but this does not guarantee the solution moves in the right direction. I use the simplest way to choose the α_k value (Tarantola, 1984): when the residual at iteration $k+1$ is bigger than the residual at iteration k, I halve the step length value. But there is an offset residual and a traveltime residual varying independently in the traveltime-offset inversion. Usually I check that at least one of them fulfils the condition.

In some examples, different sets of α values can lead towards the right solution.

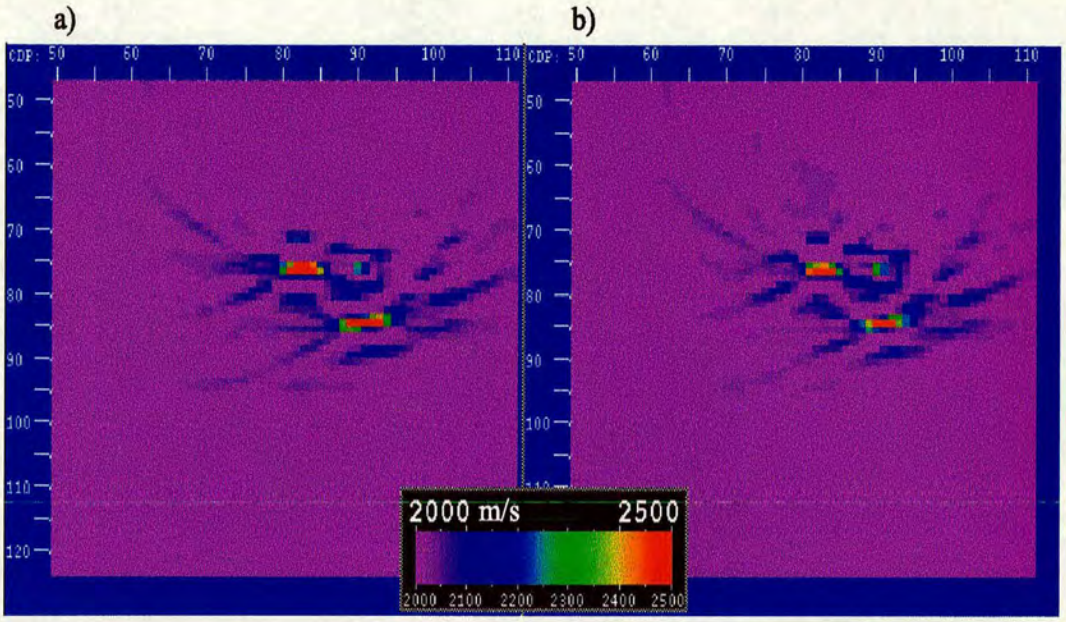


Figure 6.12: Traveltime-offset inversion using two different sets of α values

This depends on the stability of the inversion method, and on how far the current velocity model is from the real model. Figure 6.12 shows the traveltime-offset inversion using two different sets of α values. Both point-diffractors are correctly placed, and the final traveltime and offset residuals are zero in both cases; see Figures 6.13 and 6.14 for the residuals. If the convergence of the inversion method is stable, the solution oscillates around the minimum when an inappropriate step length value is used.

I now consider two additional starting models, model "a" and model "b". They are represented in Figure 6.15 by the dotted lines. The outputs from the traveltime-offset inversion indicate that the algorithm converged to a local minimum in both cases, see Figure 6.16. The position of the diffractors and the velocity above them have been adjusted to compensate for the offset and traveltime delays. The residuals are zero after 22 and 15 iterations, respectively, but the increment in the background velocity values misplaces the point-diffractors.

A way to avoid getting stuck in a local minimum has been described by several

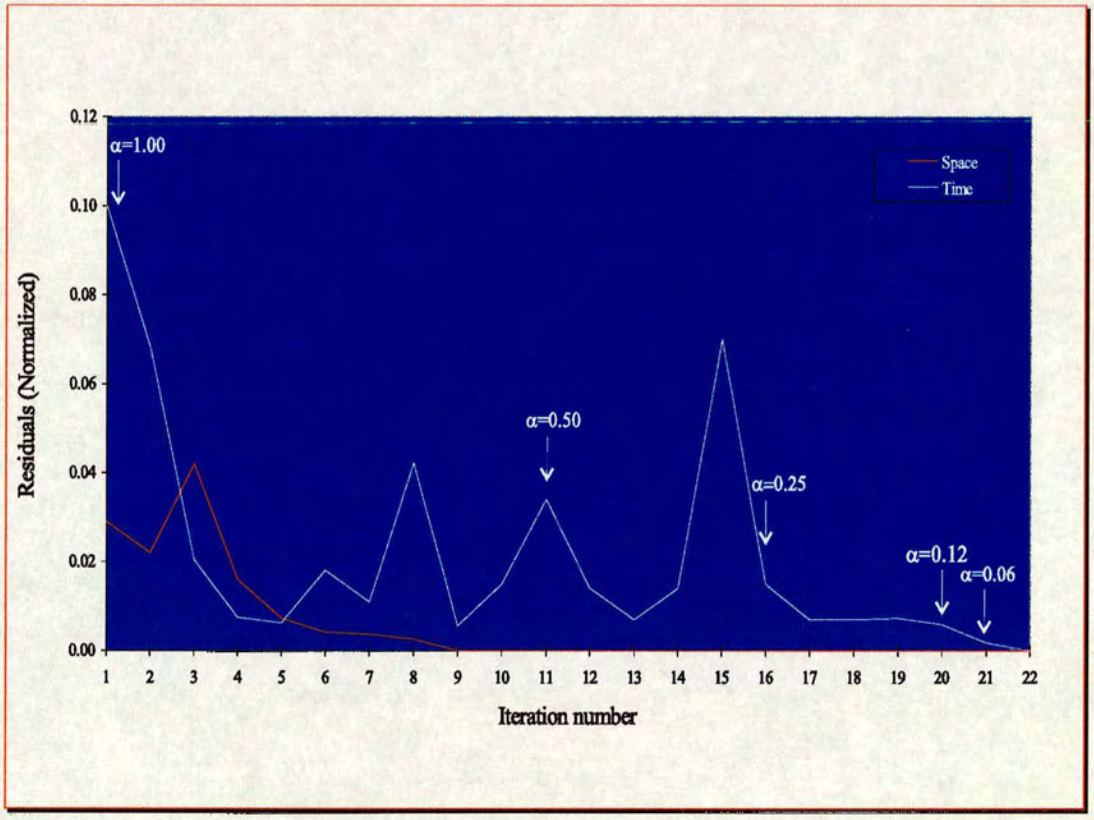


Figure 6.13: Residuals for the first set of α values.

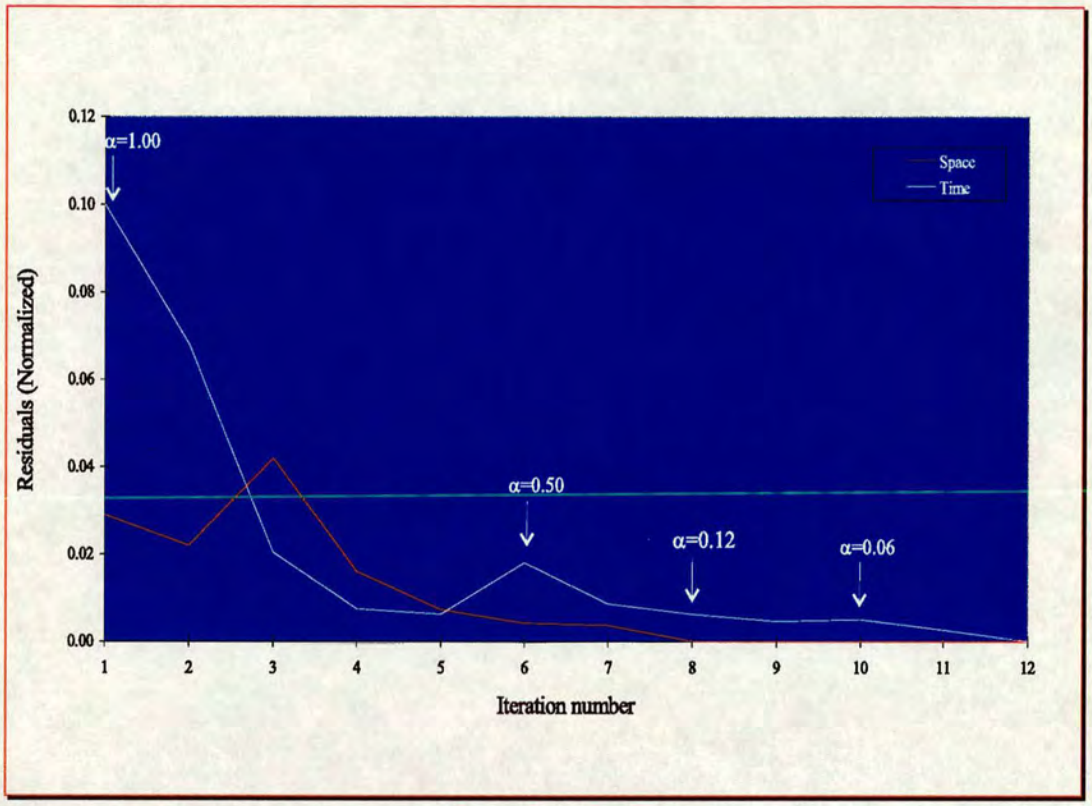


Figure 6.14: Residuals for the second set of α values.

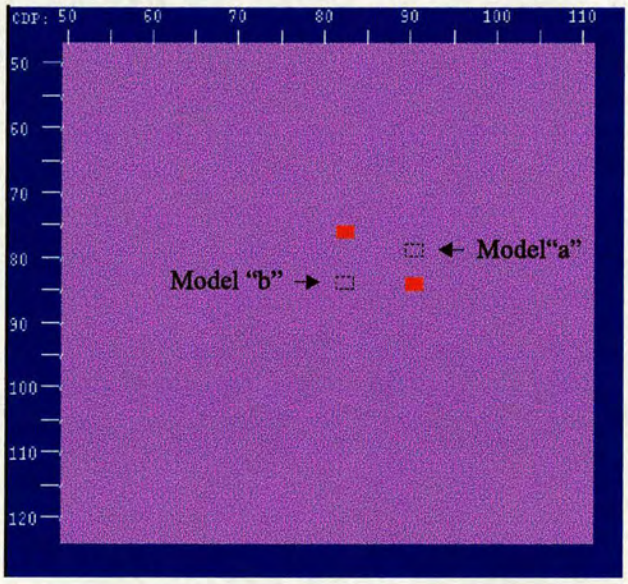


Figure 6.15: Starting models for the inversion of two diffractors laterally shifted.

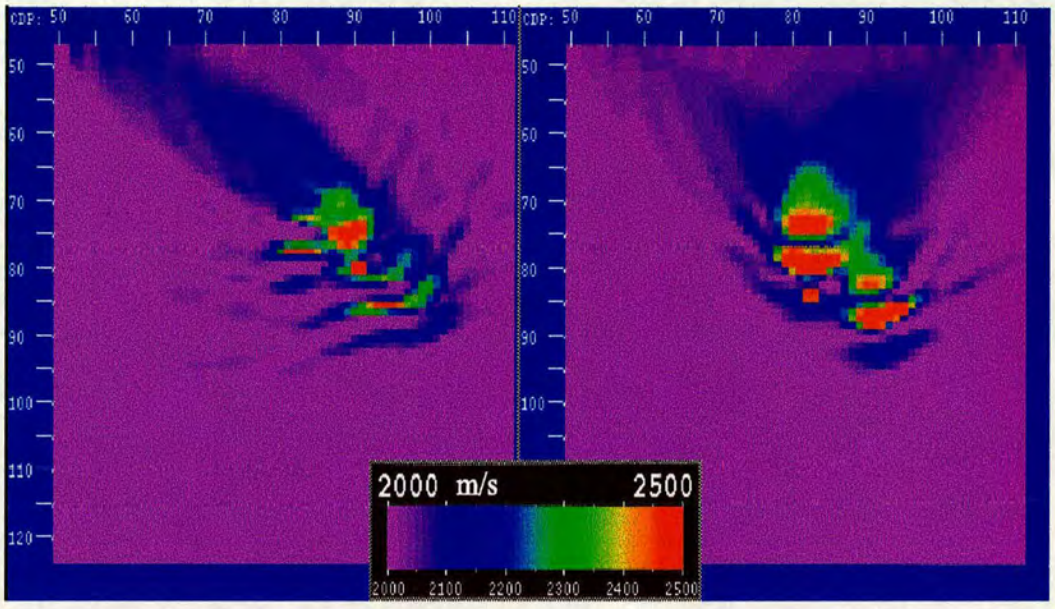


Figure 6.16: The inversion of the shifted diffractors. a) Using model "a" as starting model. b) Using model "b" as starting model.

authors. I can mention among them: Kolb et al. (1986), Kolb and Canadas (1986), Mora (1987a), Pica et al. (1990), Sun and McMechan (1992), Xu et al. (1995), and Liao and McMechan (1996). They argue that convergence to a global minimum can be achieved when the initial model produces calculated shot gathers that differ from the observed shot gathers in less than a half-period. So they low-pass filter the seismic data at the first iteration, and move up the high-cut corner of the filter as the number of iterations increases. Each frequency band satisfies the half-wavelength criterion.

I apply the above procedure to the previous example, starting with the initial model "b". Figure 6.17 displays the resulting velocity models as the frequency band increases. After 10 iterations using the first frequency band the dominant diffractor is placed in the right position. Figures 6.18a and 6.18b show the observed and the calculated shot gathers and their respective frequency spectra. There is a very good match between the original and the computed shot gathers, but the

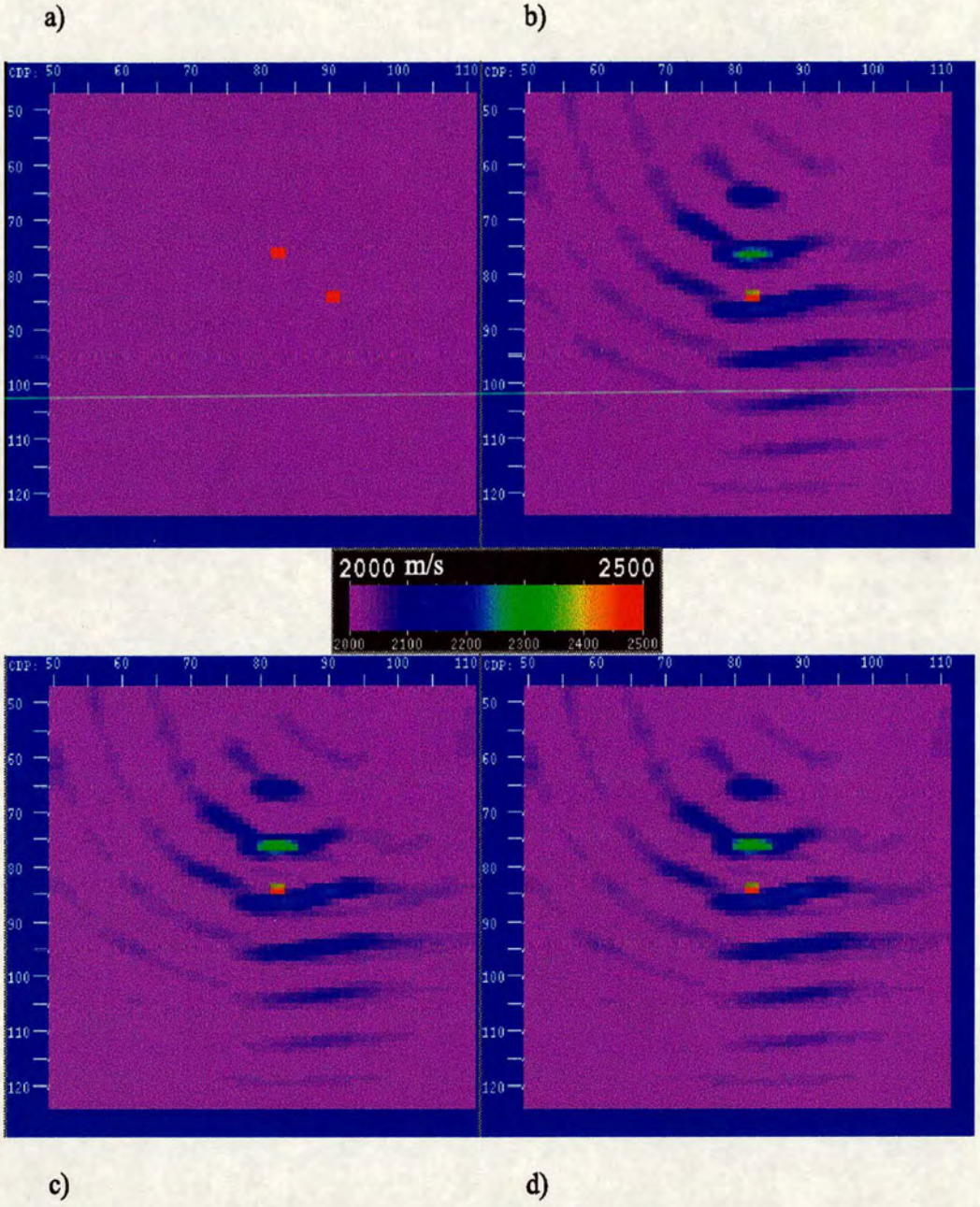


Figure 6.17: Inversion using the frequency windowing technique. a) Original model. b) Inversion after applying the first band-pass filter: high-cut corner at 80 Hz. c) Inversion using b) as starting model and the high-cut corner at 180 Hz. d) Inversion using c) as initial model and the high-cut corner at 300Hz.

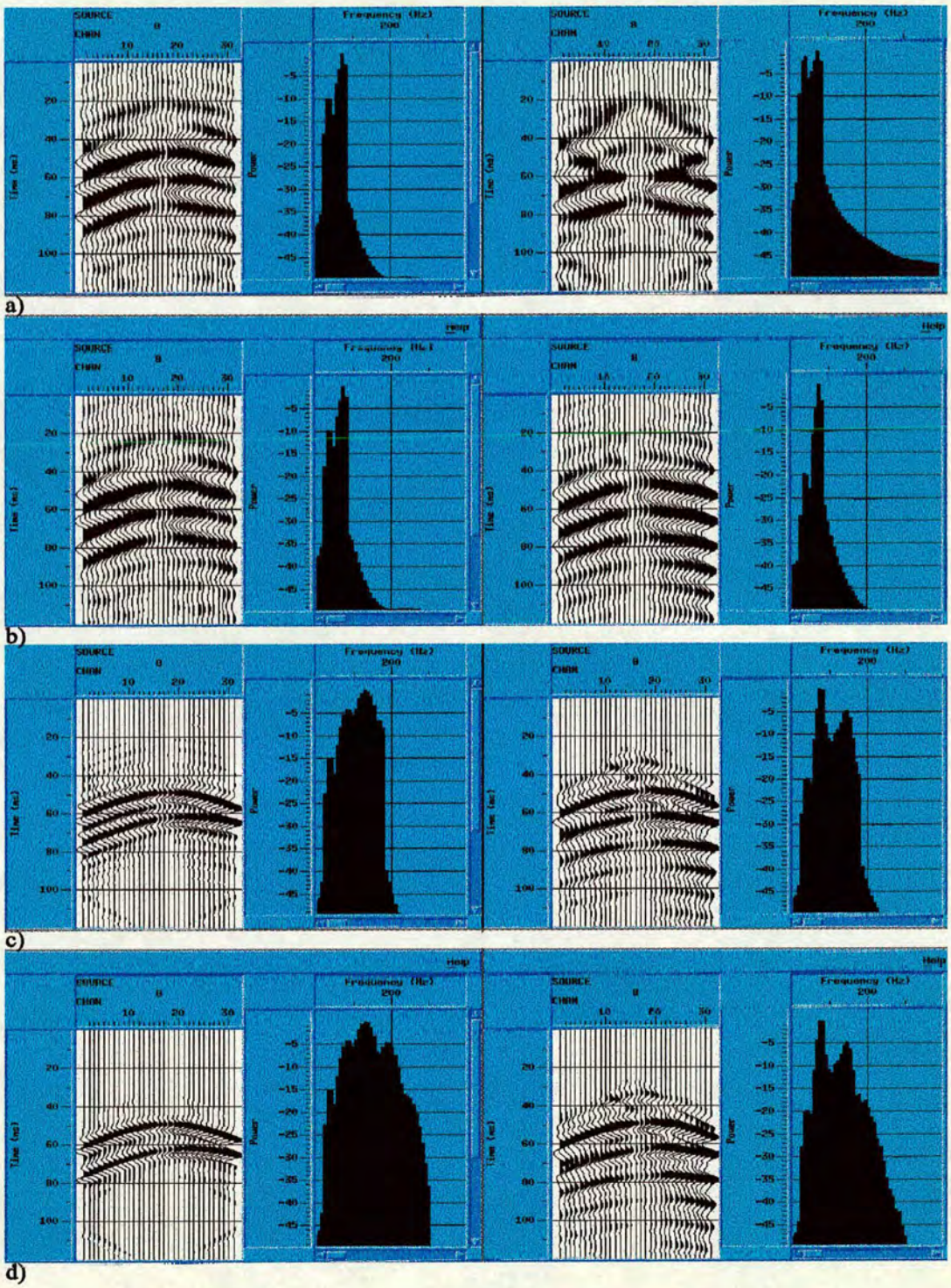


Figure 6.18: Traces and spectra from the previous example. a) Original and initial traces after applying the first filter. b), c), and d) show the results of the inversion after relaxing the frequency filter.

step length value decreases rapidly from 1.0 to 0.01 in 8 iterations. This fact does not allow the gradient to considerably increase the velocity of the reconstructed point-diffractor. For the same reason the initial point-diffractor is not removed. The successive frequency bands do not significantly alter the previous result.

6.5 Inversion of two parallel lines of point-diffractors

The model in Figure 6.19a simulates two lines of point-diffractors embedded in a constant velocity medium of 2000 m/s. The distance between the lines is 12 m, and the depth of the model is 91.5 m. Again the commencing velocity model is a single point-diffractor with the same background velocity as the original model (Figure 6.19b). The minimum velocity permitted is constrained to 2000 m/s. Figure 6.19c depicts the traveltime-offset inversion after 16 iterations. The long wavelength parts of the actual model, related to the traveltime information, are recovered.

The resolution of each interface improves after applying 10 iterations of full waveform inversion, that account for the high frequency content of the model (6.19d). For this step, I used as the initial velocity model the velocity field shown in Figure 6.19c. The resulting traces for two different shot gathers are displayed in Figure 6.20. It is observed that the calculated shot gathers correlate well with the original ones. The offset and the traveltime residuals are shown in Figure 6.21, both residuals are zero when the velocity field from the 16th iteration is used to propagate the wavefields.

6.6 A single layer with a fault

The next example illustrates the behaviour of the traveltime-offset inversion acting over a faulted layer. The original velocity field is described in Figure 6.22a. The background velocity is 2000 m/s and the velocity of the layer is 2300 m/s. The thickness of the layer is 45 m in the left hand side of the model and 97.5 m in the right hand side. 900 samples in time are recorded, and 31 receivers and 15 shots are spread along the surface. The source function continues to be a Ricker

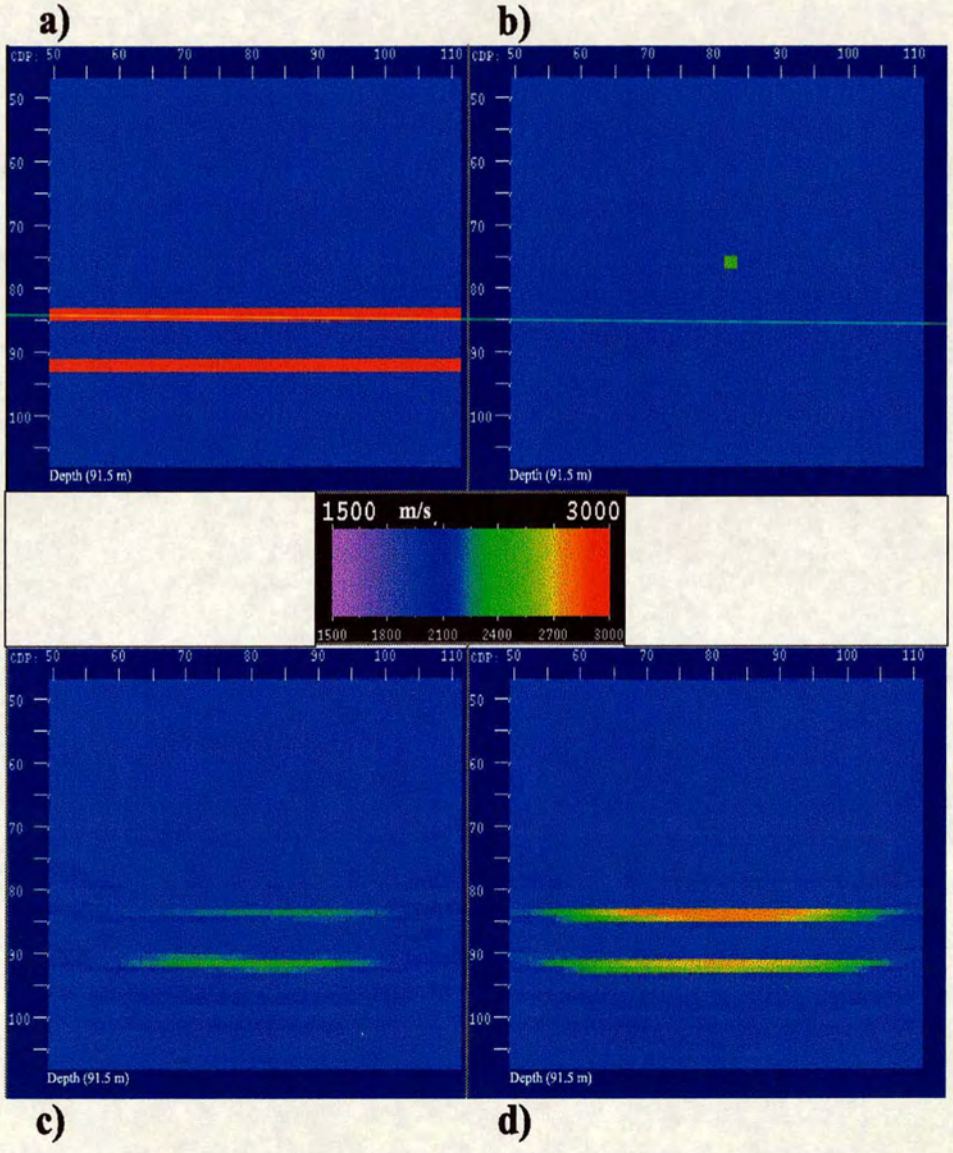


Figure 6.19: Inversion of two lines of diffractors. a) Original model. b) Starting model. c) Traveltime-offset inversion after 16 iterations. d) Full waveform inversion after 10 iterations using c) as the initial velocity model.

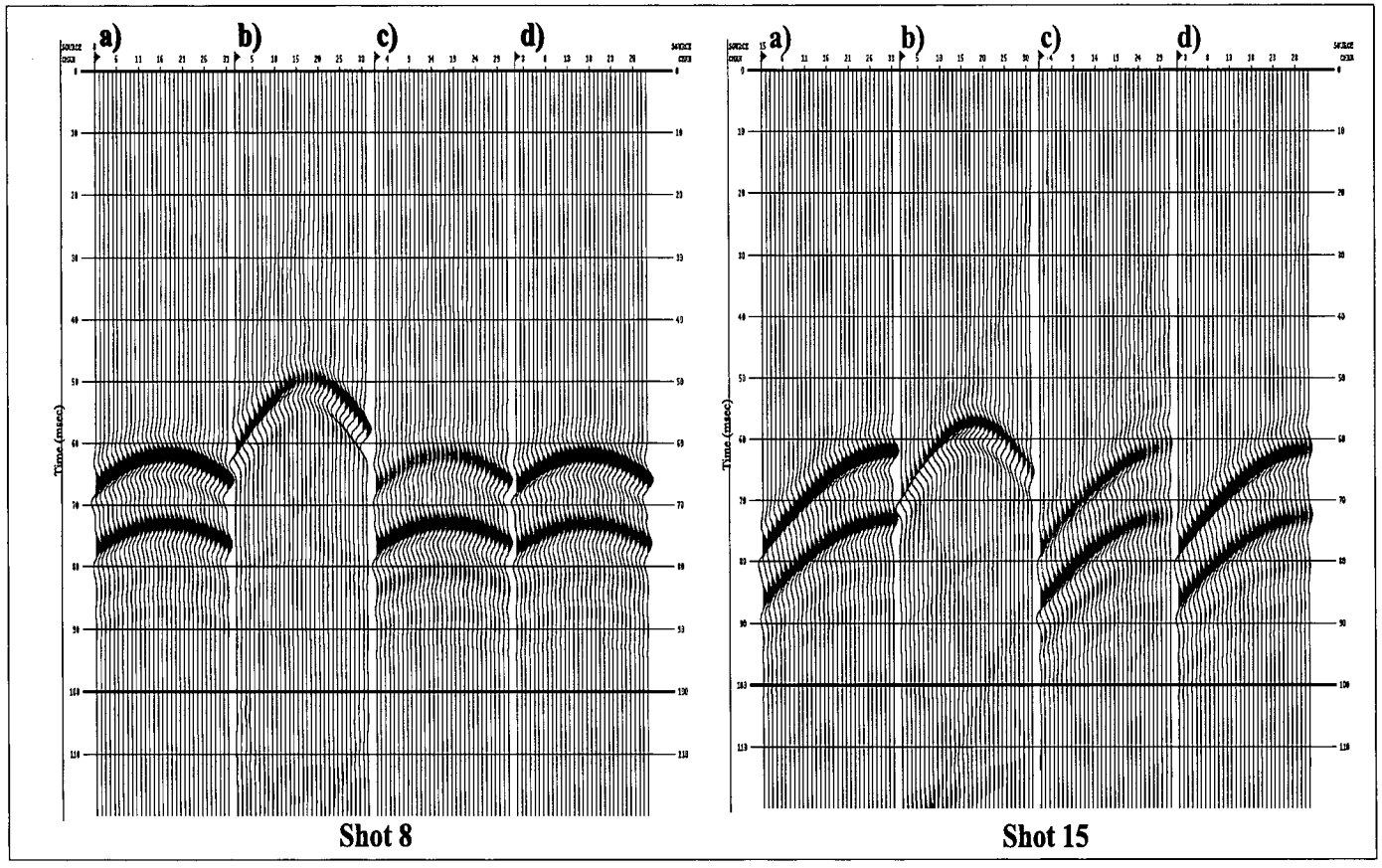


Figure 6.20: Traces from the previous example. a) Observed shot gather. b) Initial shot gather. c) Calculated shot gather after 16 iterations using traveltime-offset inversion. d) Calculated shot gather using full waveform inversion after traveltime-offset inversion.

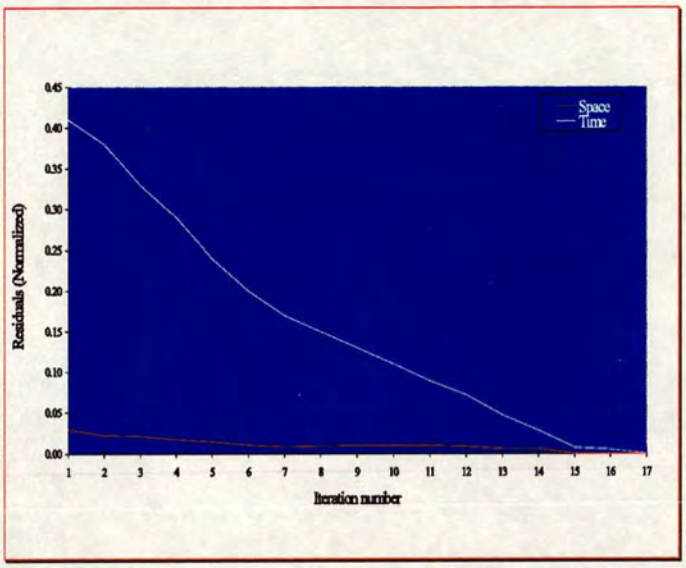


Figure 6.21: Residuals from the traveltime-offset inversion.

wavelet with 120 Hz as a peak frequency.

For the next inversions the window for the cross-correlation function has been limited in the space domain to 10 traces, i.e., the relative movement between the observed and the calculated shot gathers does not exceed 10 traces. Opposite to the full waveform inversion, the traveltime methods based on cross-correlation functions and applied to surface reflection datasets do not permit the use of constant velocity models as starting models. The cross-correlation of reflected or diffracted events is necessary to calculate the delays after muting out the direct arrivals. Therefore, I need a starting velocity model different from a constant background velocity. I use a point-diffractor because it is easy to remove it from the initial model.

Two different constraints on the minimum and maximum velocity give different responses as shown in figures 6.22b and 6.22c. In the first case the velocities are limited to the interval [2000 m/s, 3000 m/s]. After 12 iterations the output shows the presence of the interfaces, but all of them are in the wrong place and there are no appreciable differences with respect to the model after one iteration. In

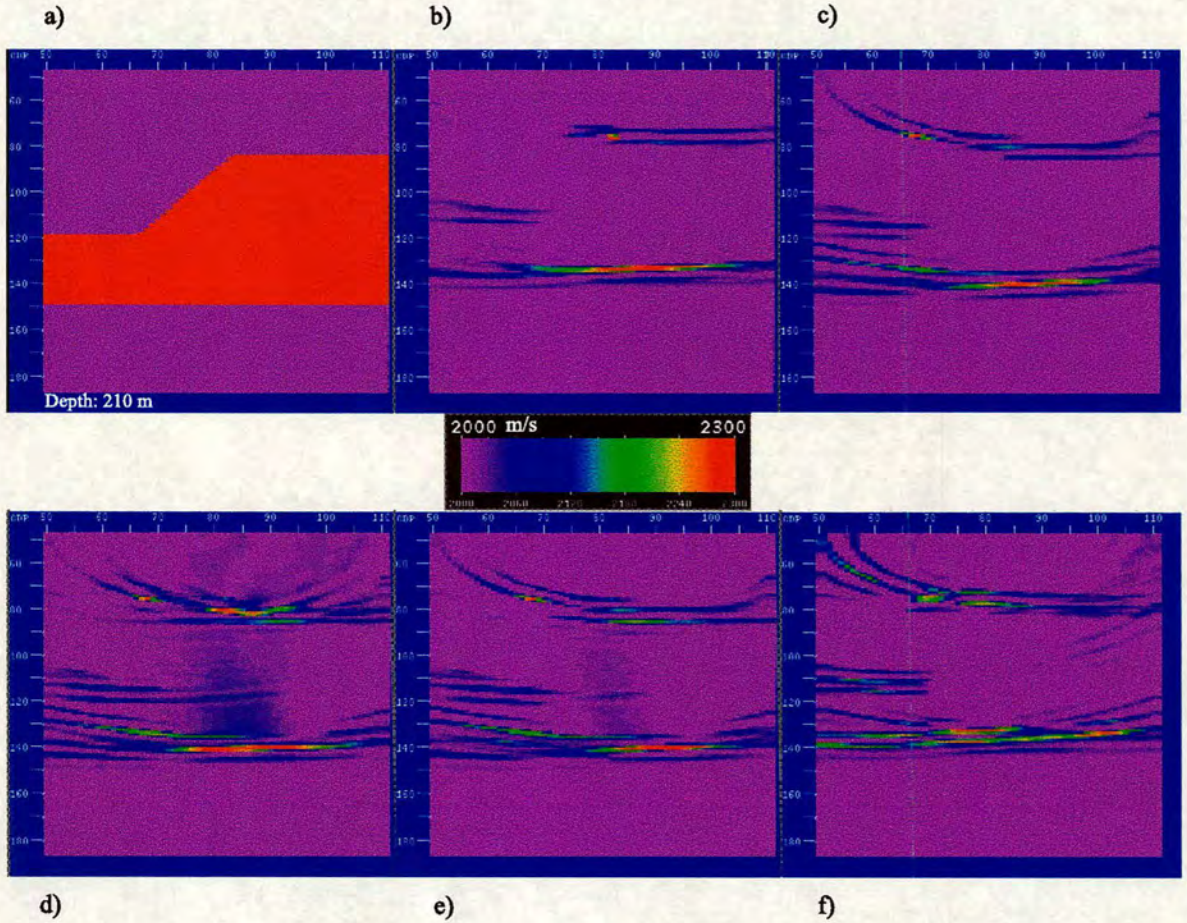


Figure 6.22: Inversion of a faulted layer. a) Original model. b) Inversion with the velocity range limited to [2000 m/s, 3000 m/s]. c) Inversion with the maximum velocity limit at 2300 m/s (19 iterations), and the maximum offset delay is equivalent to 10 traces. d) Result after 5 iterations using c) as initial guess and with the cross-correlation function considering only the last 250 samples in time. e) same as c) with the velocity of the first 78 nodes fixed. f) Inversion after 15 iterations using 5 traces as the maximum offset delay.

the second case, the maximum velocity limit is moved down to 2300 m/s. After 19 iterations the first interface is at the right place. The upper left-hand side interface is close to the correct position, but the deeper reflector is misplaced since the velocity above it did not change. The artifact in the upper layer seems to be generated by the corner of the fault.

Now that the shallower reflector has been adjusted, I try to relocate the deeper one. To do that I changed the cross-correlation window just to allow the latest event in the shot gathers to dominate the inversion. Only the last 250 samples in time contribute to the cross-correlation. Using Figure 6.22c as an initial model, the result after 5 additional iterations is displayed in Figure 6.22d. The velocity values start changing very slowly, but the velocity in the upper layer changes as well.

In order to avoid this happening, the velocity values above the node 78 in depth are kept fixed and equal to the velocities in the starting model (Figure 6.22c). The result, shown in Figure 6.22e, seems to indicate that the inversion is generating the right answer, but the convergence is very slow. Another parameter that can be changed before starting iterating is the maximum offset delay. Figure 6.22f depicts the traveltime-offset inversion after 15 iterations when this value is set up to 5. This result is comparable to the inversion using $\Delta\sigma_{max} = 10$, represented in Figure 6.22c.

The traveltime-offset inversion for the same model is redone using the frequency windowing technique to recover the lower wavenumber features first. A high-cut filter at 80 Hz is applied. The result, after 12 iterations, shows an increment in the velocity values in the layer (Figure 6.23b). However, when the high-cut point is relaxed (150 Hz) the inversion goes in the opposite direction. After 7 iterations the deeper interface goes up, and the velocity is reduced to adjust the traveltime. Figures 6.23c, 6.23d, and 6.23e display the results after applying high-cut filters of 60 Hz, 80 Hz, and 120 Hz progressively as iterations increase. The response of the inversion when the first band has a lower high-cut

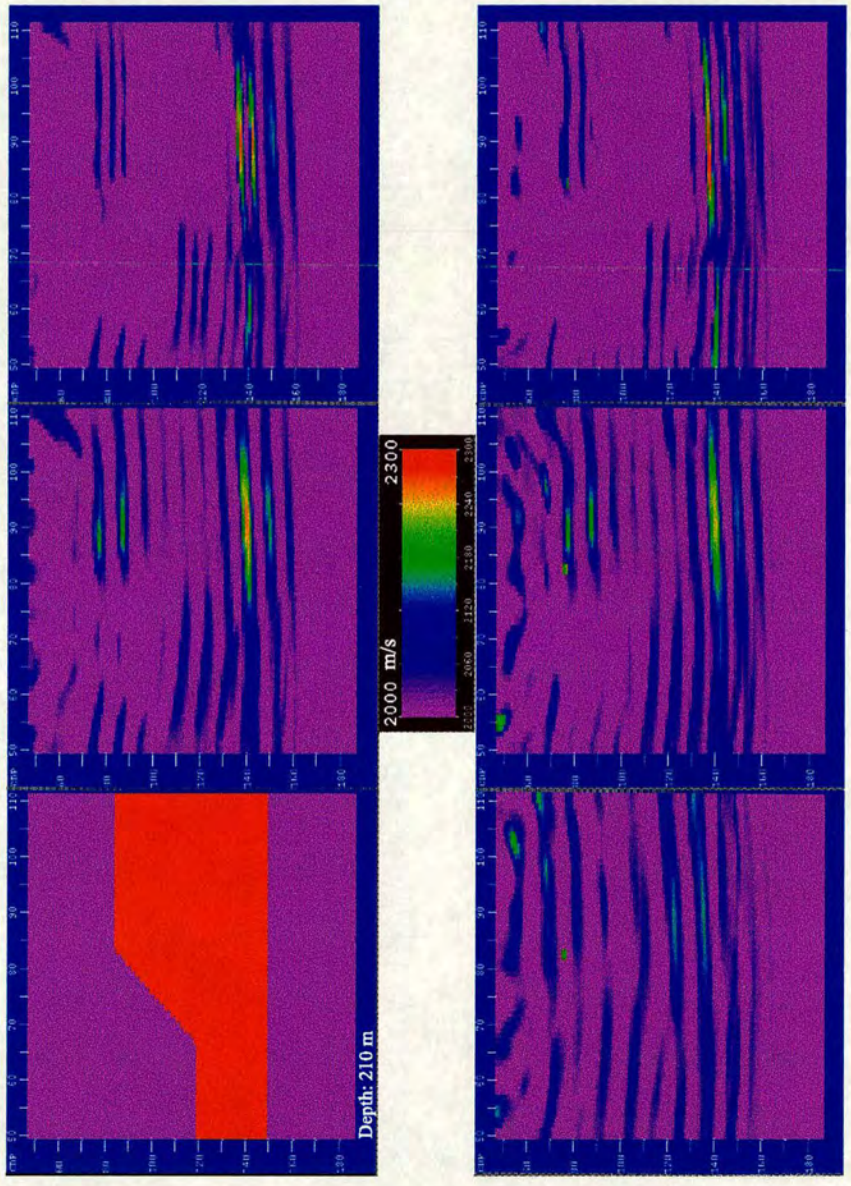


Figure 6.23: Inversion of the faulted layer using the filtering technique. a) Original velocity model. b) High-cut corner at 80 Hz. c) High-cut corner at 150 Hz. d) A second attempt starting with the high-cut corner at 60 Hz. e) High-cut corner at 80 Hz. f) High-cut corner at 120 Hz.

corner is similar to the previous result.

Another attempt to recover the velocity of the faulted layer consists of alternating the standard traveltime inversion (one-dimensional cross-correlation) and the traveltime-offset inversion (bi-dimensional cross-correlation). The first inversion on its own is displayed in Figure 6.24b after 15 iterations; the second one is shown in Figure 6.24c after 19 iterations, and the combination of both is shown in Figure 6.24d. For the last inversion a total of 10 iterations is performed and the result is comparable to Figure 6.22d. It can be seen, in figures 6.24b and 6.24c, that the output from the standard traveltime inversion differs markedly from the result using the traveltime-offset inversion. The first method gets stuck in the point-diffractor and creates artifacts around it. The second one is more stable and, at least, detects the presence of the interfaces.

The next example is the same structure, but thinner: 30 m depth on the left hand side and 54 m depth on the opposite side (see Figure 6.25a). Figures 6.25b and 6.25c illustrate the traveltime-offset inversion at iterations 12 and 27, respectively. After 12 iterations, the velocity values above the node 78 are fixed. The definition of the structure improves, but the inversion seems to be affected by the poor illumination near the borders. Further iterations using the full waveform inversion smooth the output, see Figure 6.25d.

The traveltime-offset inversion using the windowing filters is displayed in Figure 6.26. This time, the velocity above the node 78 can vary. The lower frequency inversions introduce some artifacts in the upper layer, that can be related to an imperfect mute of the first arrivals. As the velocity in the upper layer is higher than the original background velocity, the last interface is placed deeper than the real one. A high frequency pattern, parallel to the interfaces, is observed in this figure. Again, the full waveform inversion smooths the output (Figure 6.26f). Even though the interfaces are located at the correct depth and well defined, the velocity of the faulted layer is balanced with higher velocities in the upper layer.

The next step is to extend the model laterally to see if a better image of the

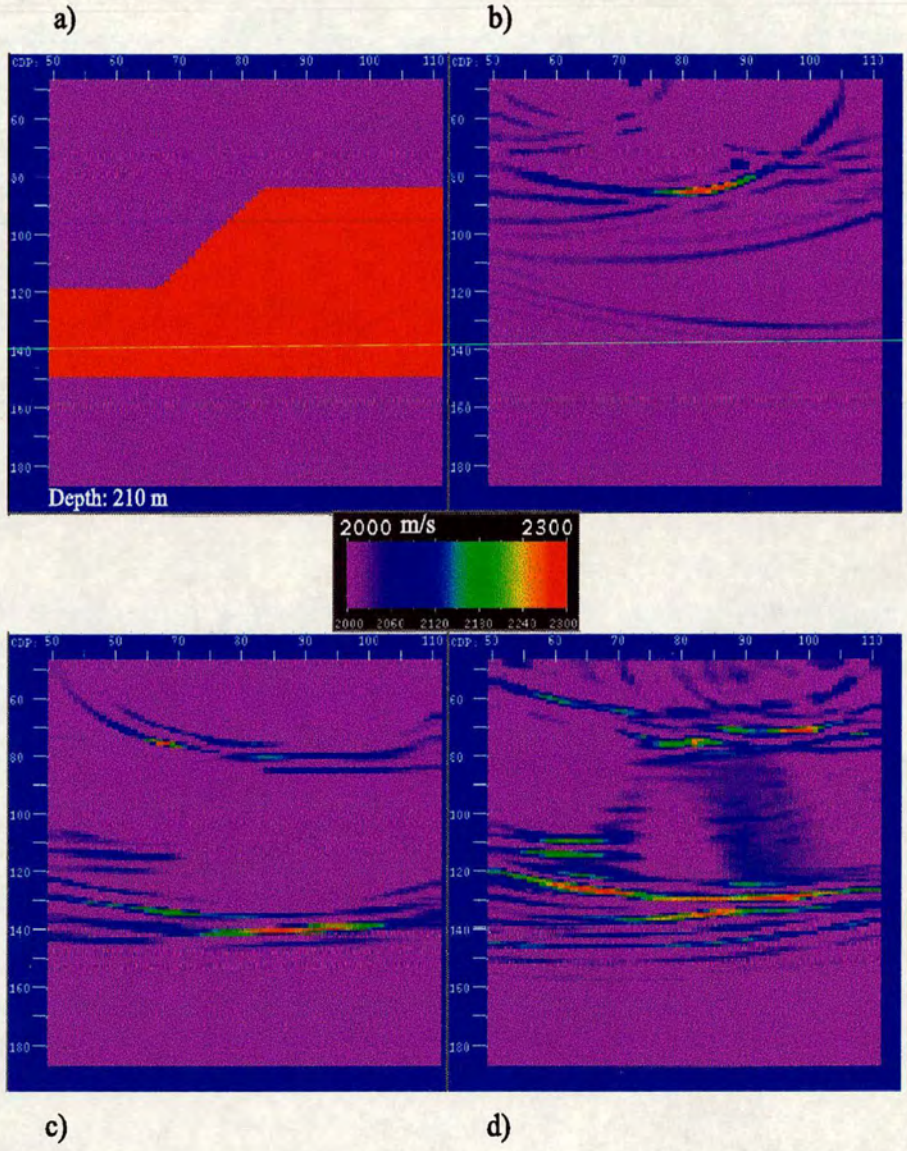


Figure 6.24: Alternating the standard traveltime inversion and the traveltime-offset inversion. a) Original model. b) Standard traveltime inversion after 15 iterations. c) Traveltime-offset inversion after 19 iterations. d) The alternated inversion after 10 iterations.

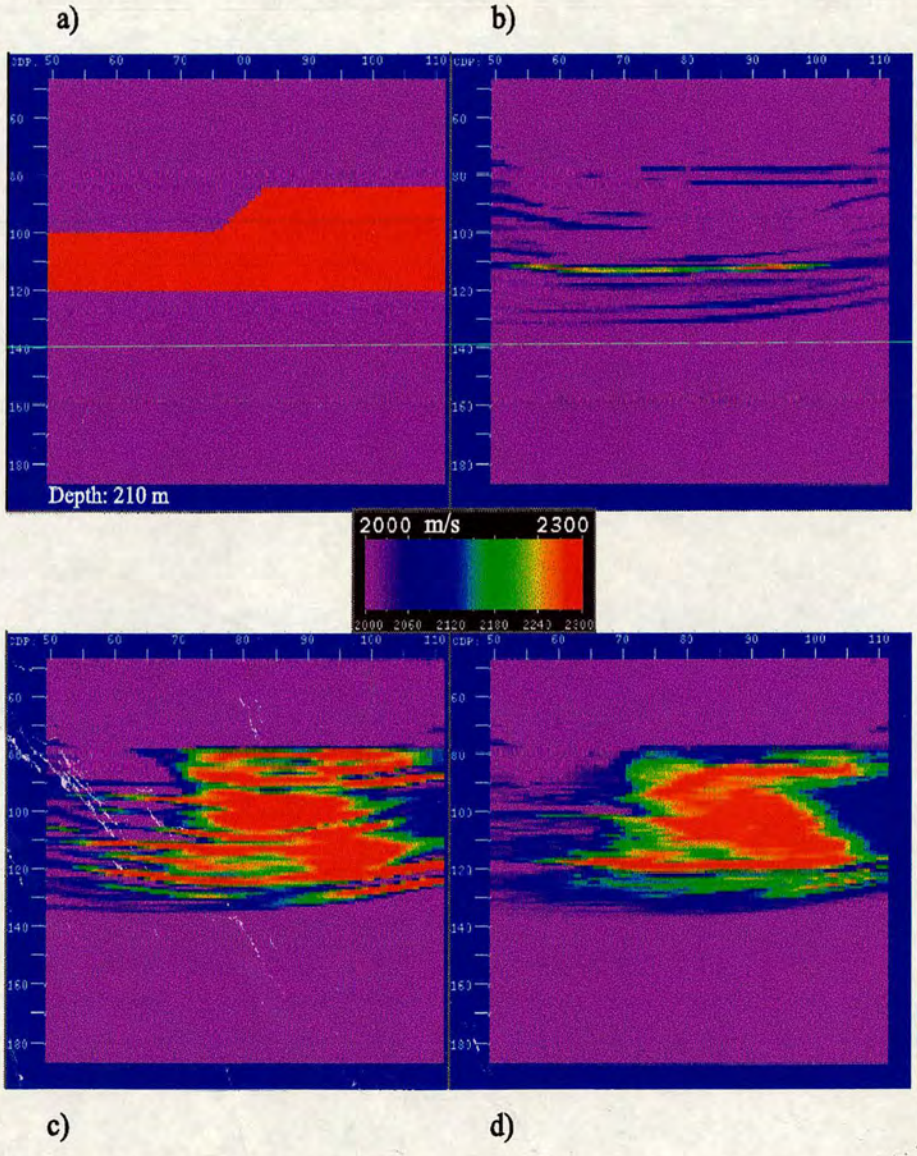


Figure 6.25: Inversion of a thin layer. a) Original velocity model. b) Inversion after 12 iterations. c) Inversion using b) as initial model but keeping the velocity fixed above the node 78 (15 iterations). d) Full waveform inversion using c) as initial guess.

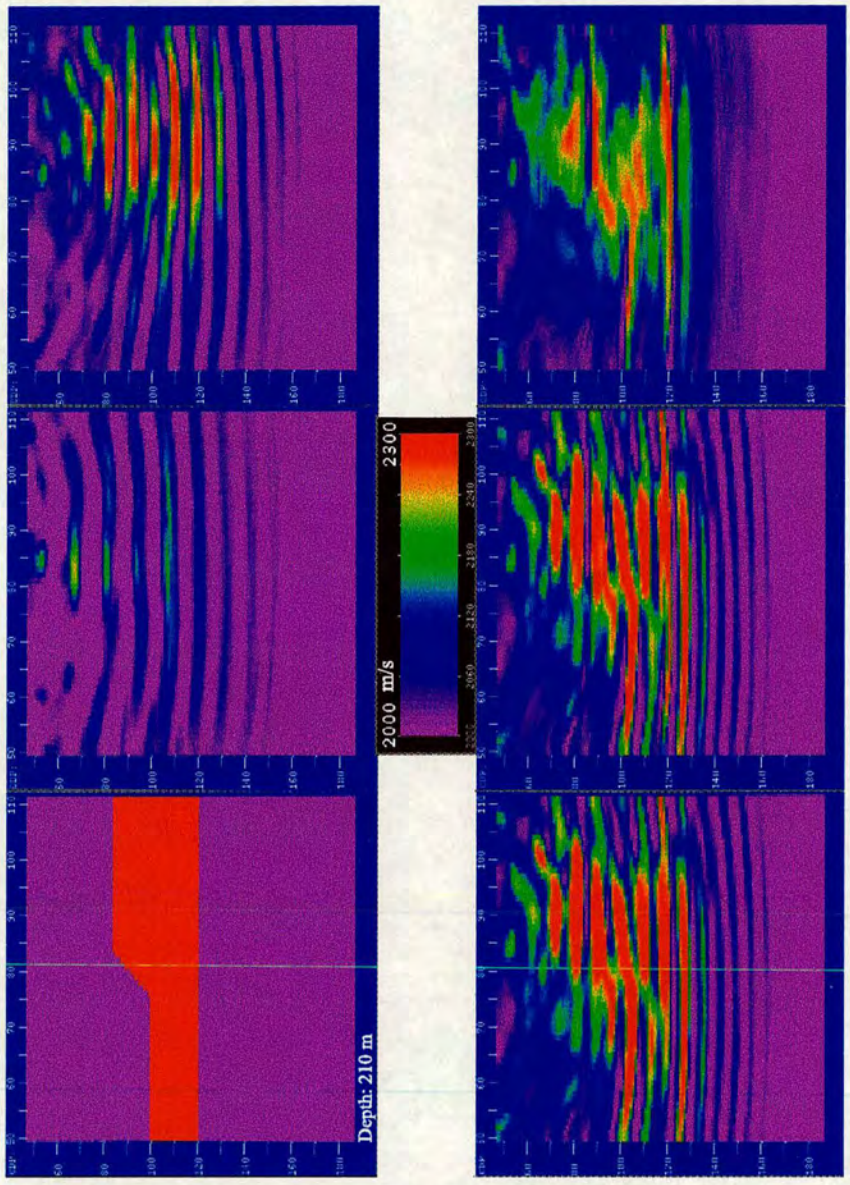


Figure 6.26: Inversion of a thin layer using the frequency windowing technique. a) Original model. b) High-cut corner at 60 Hz (6 iterations). c) High-cut corner at 80 Hz (7 iterations). d) High-cut corner at 100 Hz (4 iterations). e) Full-frequency inversion (5 iterations). f) Full waveform inversion using e) as starting model.

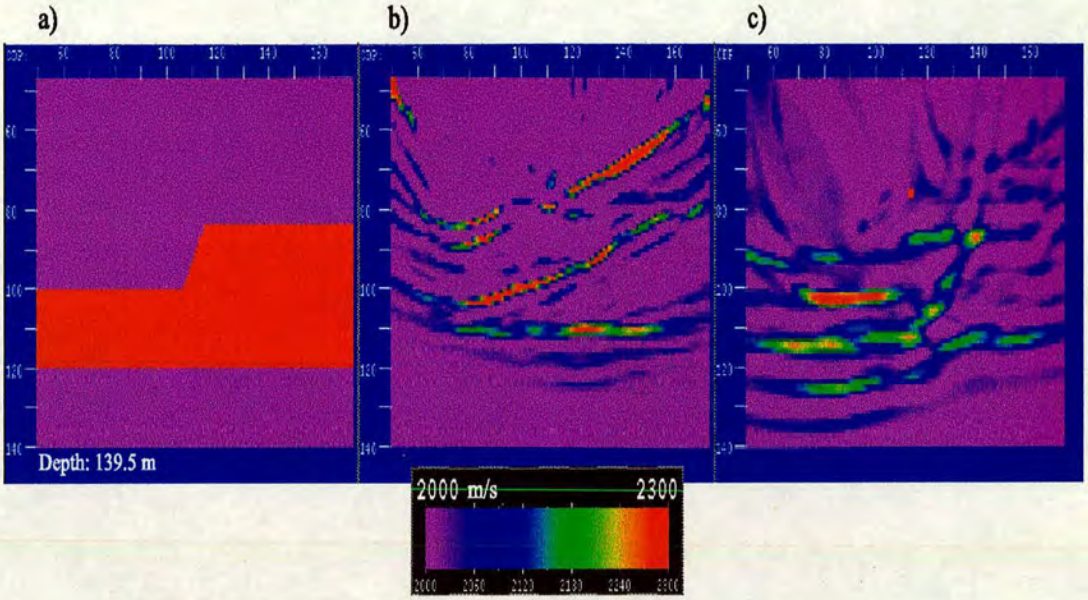


Figure 6.27: Inversion of an extended thin layer using two different Ricker wavelets. a) Original velocity model. b) Peak frequency: 120 Hz, and c) Peak frequency: 60 Hz.

structure is obtained. To do that, 31 nodes are added to both lateral boundaries in the previous example but, due to computational reasons, only 94 nodes in depth are considered. The 15 shots and 31 receivers are located at the same coordinates as before. The new area is covered with 31 additional receivers, regularly spread on the surface. Figure 6.27b shows that the traveltime-offset inversion fails to position the reflectors after 18 iterations. The result improves slightly if the peak frequency of the Ricker wavelet is changed from 120 Hz to 60 Hz (Figure 6.27c).

Applying the same sequence of filters used in Figure 6.26 enhances the images when using the Ricker wavelet with peak frequency of 120 Hz. Figure 6.28 depicts the results. The faulted reflector is placed correctly, but there are still some artifacts in the upper layer. Since the velocity of the central layer is not well recovered the last interface is deeper than the original interface.

The sequence in Figure 6.29 describes the traveltime-offset inversion of the previous extended model starting with the velocity field shown in Figure 6.29b.

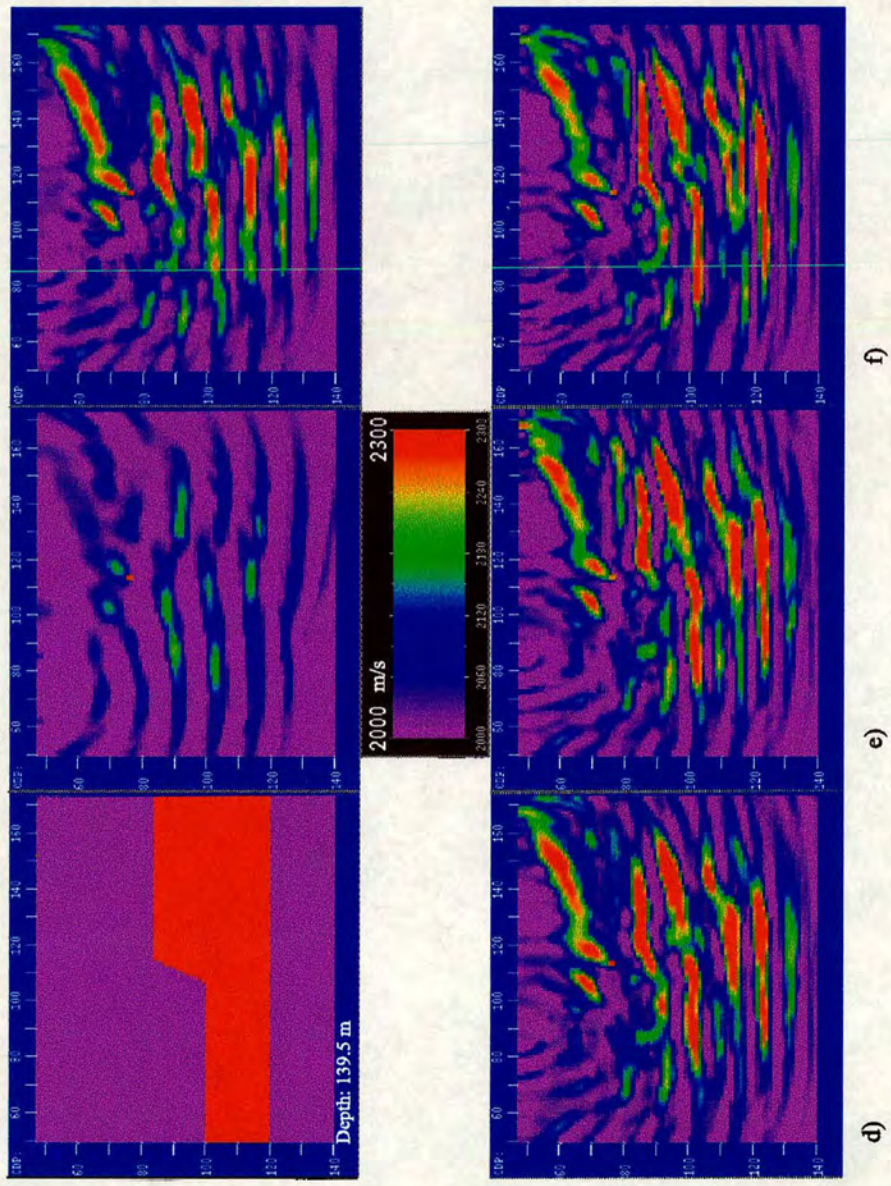


Figure 6.28: Inversion of the extended thin layer using filters. a) Original model. b) High-cut corner at 60 Hz. c) High-cut corner at 80 Hz. d) High-cut corner at 100 Hz. e) High-cut corner at 160 Hz. f) Full frequency content.

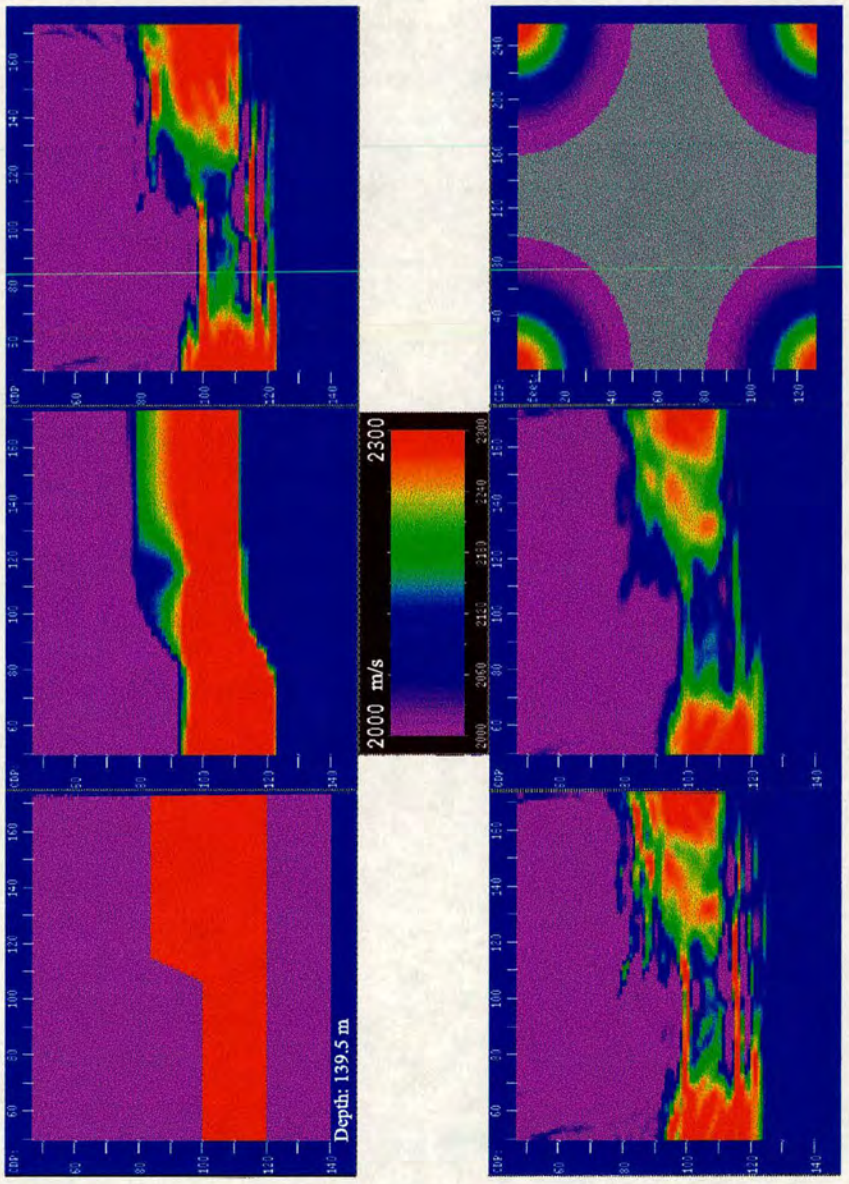


Figure 6.29: Inversion using stacking velocities and smoothing. a) Original velocity model. b) Starting model (stacking velocities converted to interval velocities). c) Traveltime-offset inversion after 8 iterations. d) The smoothing operator shown in f) is applied after each iteration. e) The velocity field in d) after applying the smoothing operator (iteration number 8). f) Gaussian filter in the (k_x, k_z) domain.

This initial velocity model represents the interval velocities converted from a set of stacking velocities. The stacking velocities were picked and converted using "Pro-MAX". Again, the result after 8 iterations shows the inversion succeeds in locating the interfaces, but fails to recover the velocity of the central layer. Smoothing the output at each iteration, by filtering it in the wavenumber domain, does not markedly improve the result. Figures 6.29d, 6.29e, and 6.29f show, respectively, the inversion using the mentioned bi-dimensional filter after 8 iterations, the same velocity field after smoothing, and the Gaussian filter in the (k_x, k_z) domain.

The inversion of a structure having characteristic wavelengths comparable to one wavelength and half wavelength of the peak frequency in the source is shown in figure 6.30. The depths of the model are 9 m and 18 m in the left and right hand side, respectively. The peak frequency of the Ricker wavelet is 120 Hz (wavelength = 16.7 m). For the inversion in Figure 6.31b the maximum offset delay permitted is 3 traces, while for Figure 6.30b it is 6 traces.

The result using $\Delta\sigma_{max} = 3$ shows a clearer image of the structure than the result from $\Delta\sigma_{max} = 6$. The inversion using 6 traces as the maximum offset delay diverges. However, if the band-pass filter procedure is used, the inversion is more stable and the interfaces are defined, see figures 6.30c, 6.30d, and 6.30e. Applying full waveform inversion after traveltime-offset inversion improves the image, but still the velocity field is not well recovered. An $\alpha - trimmed$ smoothing operator (Gersztenkorn and Scales, 1988) has been applied to the inversions shown in Figure 6.30.

It can be observed that there is a connection between the results from the inversions and the frequency content of the wavelet. The traveltime-offset and full waveform inversions are unable to reconstruct features with characteristic lengths of the order of half the wavelet length in a few iterations. Figure 6.32 depicts the full waveform inversion using 60 Hz as a peak frequency. The part of the structure having a 9 m thickness (approximately $\lambda_\nu/4$) is well defined, but the thicker part is not.

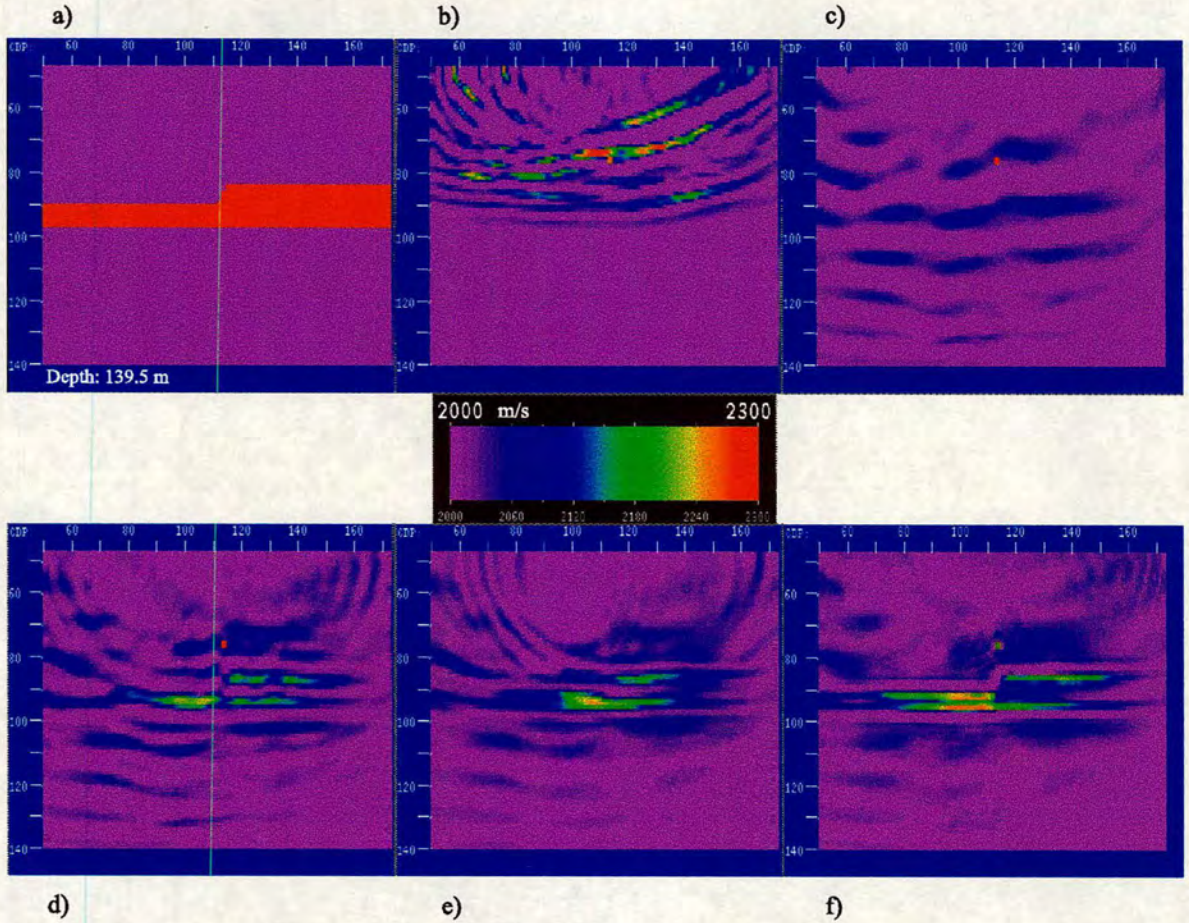


Figure 6.30: Inversion of a thin layer with $\Delta\sigma_{max} = 6$. a) Original model. b) Traveltime-offset inversion. c) Inversion using a band-pass filter: 2-5-60-65. d) Filter: 2-5-120-130. e) Full frequency content and $\alpha - trim$ smoothing. f) Full waveform inversion using d) as initial model.

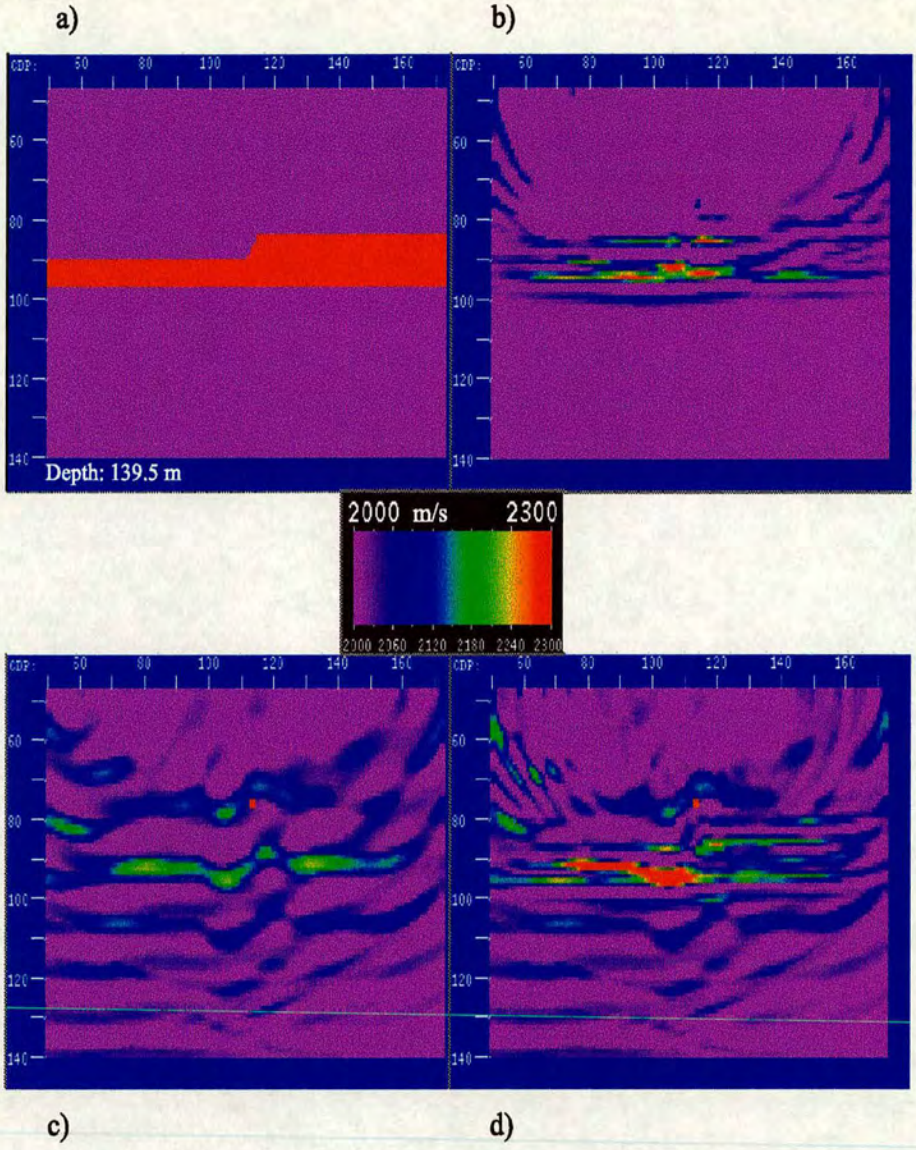


Figure 6.31: Inversion of a thin layer with $\Delta\sigma_{max} = 3$. a) Original model. b) Traveltime-offset inversion. c) Filter: 2-5-6-65. d) Full frequency content.

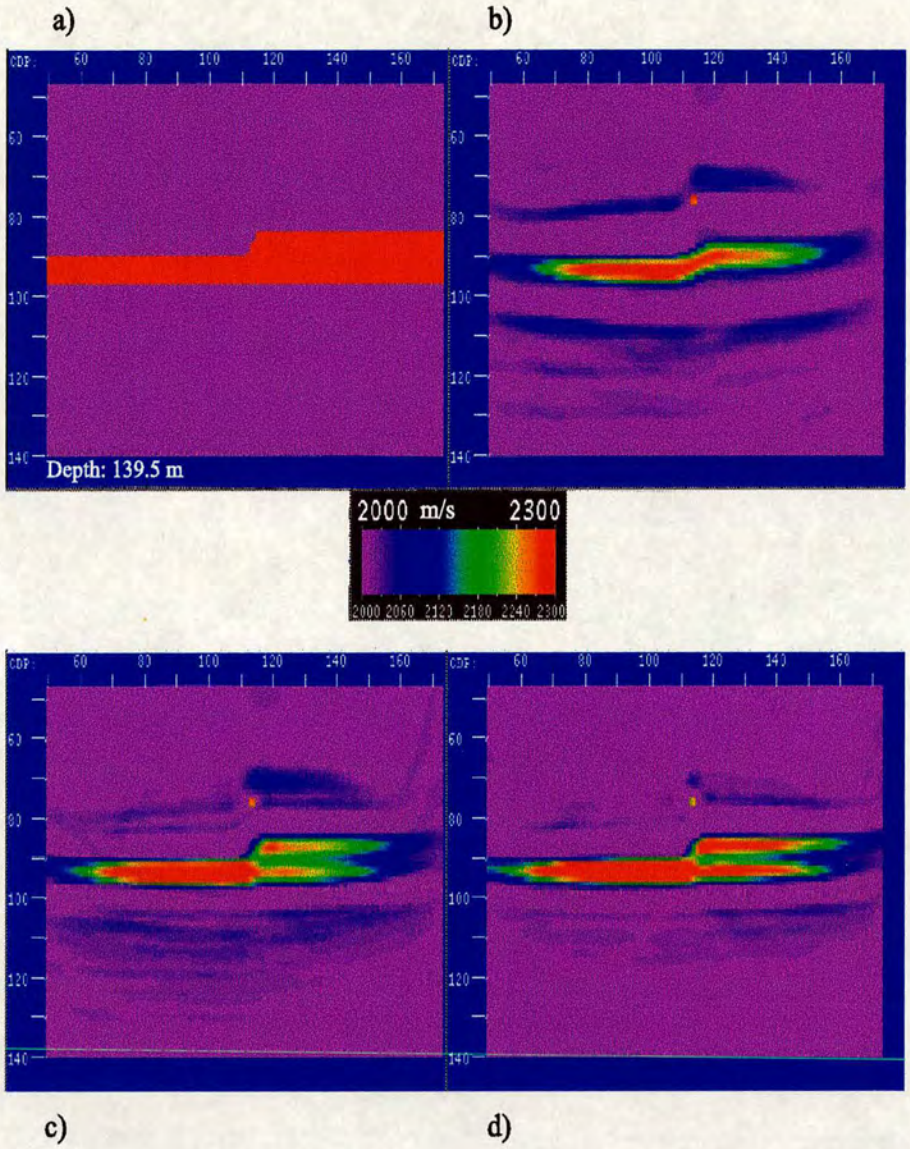


Figure 6.32: Full waveform inversion of the thin layer. a) Original model. b) Inversion after applying a high-cut filter at 60 Hz (4 iterations). c) Inversion using b) as initial guess and with the full frequency content (4 iterations). d) Full waveform inversion without filtering (14 iterations).

The convergence for the thicker part of the model is very slow, and a large number of additional iterations are needed to reconstruct it. Changing the pre-conditioning factor and calculating precisely the value of the step length may accelerate the convergence. I emphasise here that the traveltime-offset inversion recovers the very low frequency content of the velocity models by positioning correctly the dominant reflector using the traveltime information. Additionally, the proposed method maintains the original moveout between traces during back-propagation, that helps to define the shape of the interfaces.

6.7 A scheme for full waveform and traveltime-offset inversion

I mentioned in Chapter 3 that the full waveform and the traveltime-offset inversions differ, mainly, in the misfit functions and in the residuals to be back-propagated. Here I show, schematically, how these residuals are built up and their relation to the output. The wavelets are hypothetical and let's suppose that each positive peak corresponds to an event. These Ricker wavelets have been generated using "Maple" (mathematical calculation package), and the peak frequency is 120 Hz.

I now assume that one shot with one trace is recorded, see Figure 6.33a, and that the real velocity model is a point diffractor in a constant velocity medium. If the initial velocity model is similar to the original but with the point-diffractor in a different position, the functions to be back-propagated in each inversion are shown in figures 6.33c and 6.33d. As can be seen in figure 6.33c, δp has 3 positive and 3 negative peaks, so the full waveform gradient will have 3 positive and 3 negative bands which alternate as shown in Figure 6.34c. The main negative peak removes the initial diffractor, and the main positive peak creates the missing diffractor.

Each event in a trace generates one of these alternating bands and they are superimposed on the gradient. This fact explains the patterns I observed in seve-

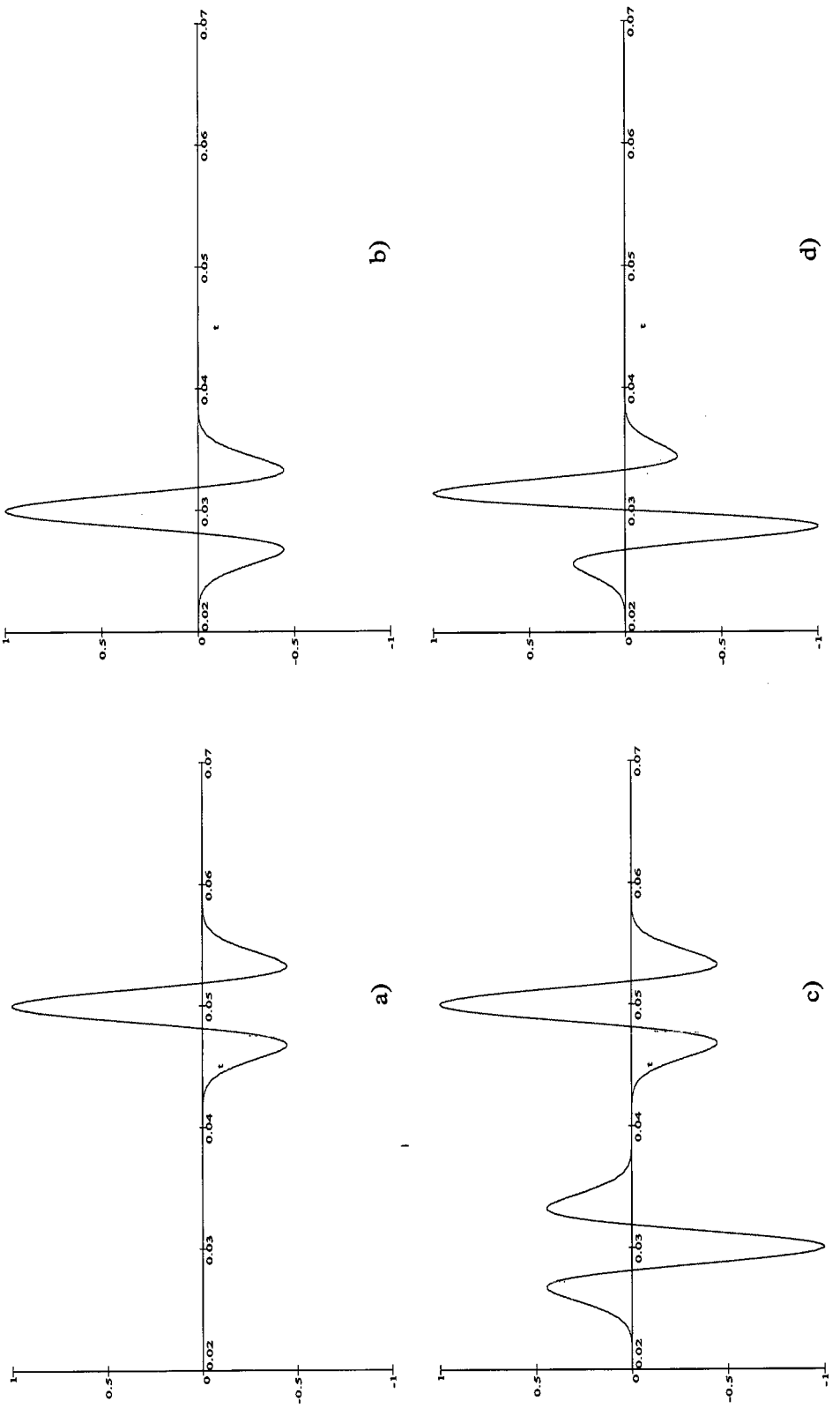


Figure 6.33: Calculating the pseudo-residuals. a) Observed trace. b) Calculated trace. c) Seismogram residual (δp) to be back-propagated. d) Traveltime residual ($\delta\tau$) to be back-propagated.

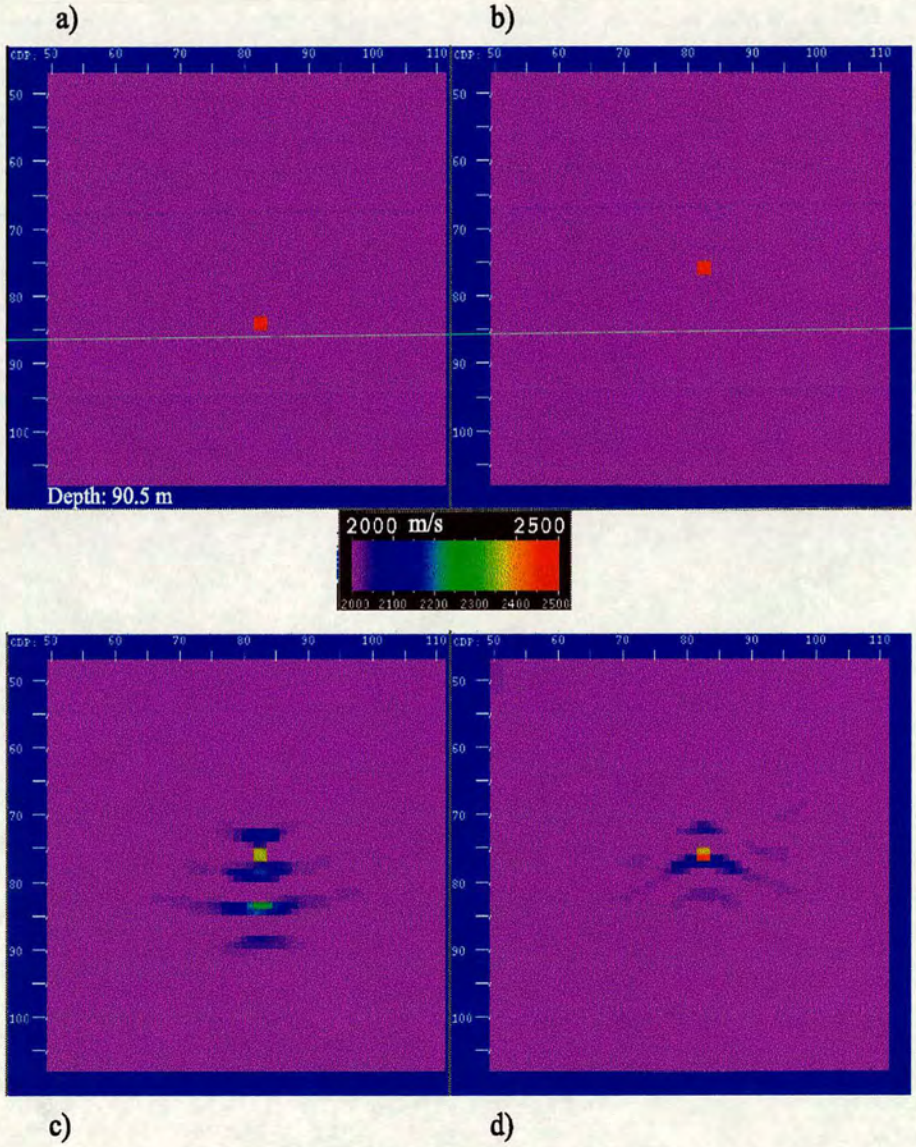


Figure 6.34: Inversion of a point-diffractor. a) Original model. b) Starting model. c) Full waveform inversion (1 iteration). d) Traveltime-offset inversion (1 iteration).

ral examples, including those in Chapter 3. The width and regularity of the pattern depend on the type of wavelet and on its frequency spectrum. There is a pattern in the case of traveltime-offset inversion as well, but it is slightly different. Figure 6.33d shows that the residual to be back-propagated has two positive and two negative peaks. The pattern in the gradient is positive, negative, positive, and negative again (see Figure 6.34d). The main negative peak removes half of the initial diffractor, while the main positive peak generates part of the missing diffractor. It seems as though, as iterations proceed, the inversion moves the point diffractor down until the delays are zero.

Figures 6.35a and 6.35b sketch the traveltime-offset and the full waveform inversions when there are two diffractors in the original model as well as in the starting model. If the shallow diffractor dominates the bi-dimensional cross-correlation function, the inversion moves the upper diffractor down in the current velocity model and adds a second diffractor a meters deeper. As the background velocity is correct, the distance a in the current velocity model is correct as well. The second diffractor in the starting model remains unaltered, because the residual to be back-propagated does not carry information about it. The same procedure applies when the deeper diffractor in the starting model dominates the cross-correlogram (Figure 6.35b). On the other hand, the full waveform inversion removes the initial diffractors, and locates the missing diffractors in the right place (if the background velocity is correct).

Figure 6.35c illustrates the case when the perturbation in the medium is a continuous layer of thickness a and constant velocity V_2. In both inversion methods, the real shot gathers carry information about two reflection events separated in time. Then, the traveltime-offset inversion moves the dominant reflector down in the starting model and adds a second reflector b meters apart, where b is related to the background velocity V_1 and not to V_2. The lengths d and e depend on the frequency spectrum of the wavelet and play an important role when trying to obtain the "blocky" effect. This scheme explains the traveltime-offset inversion

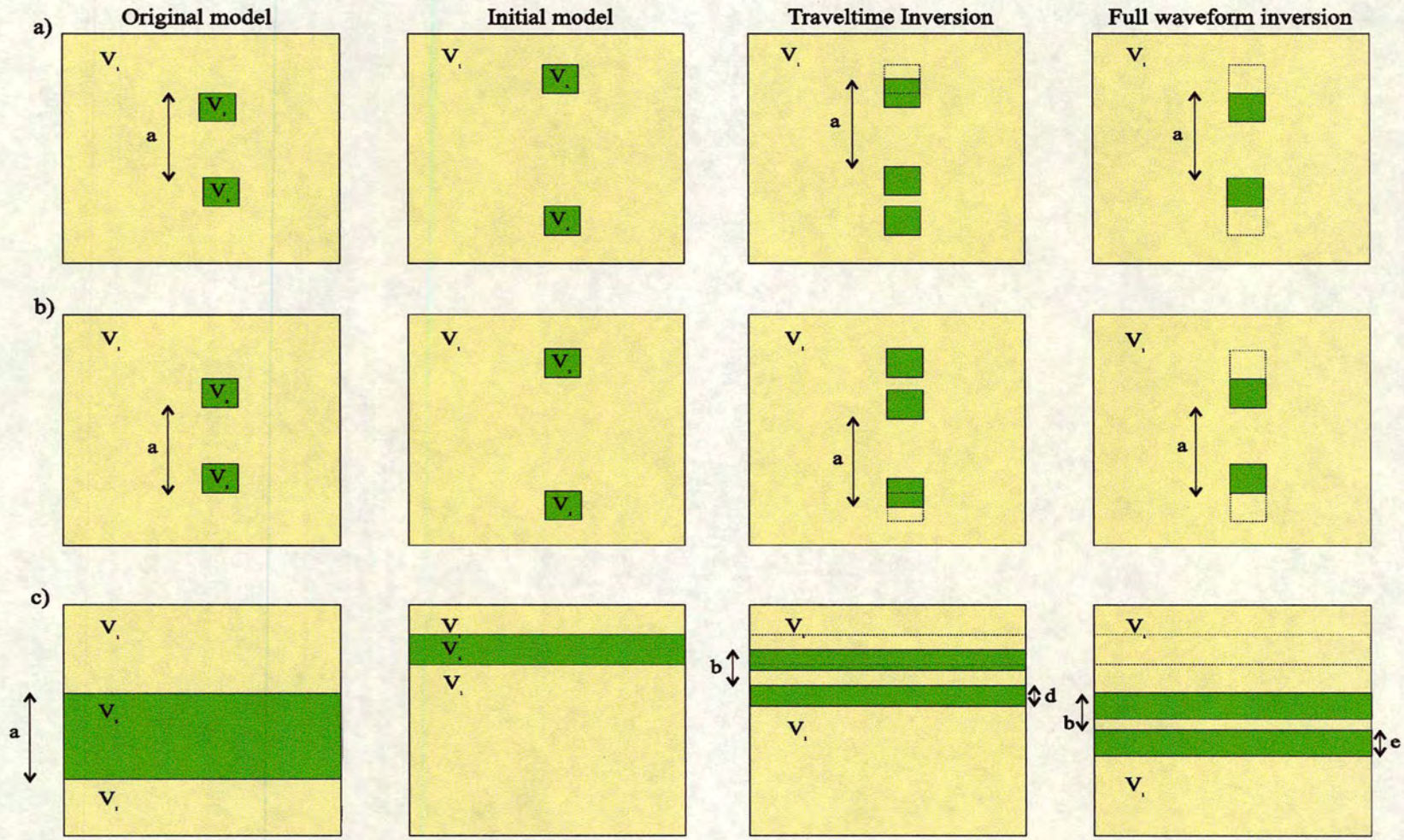


Figure 6.35: Scheme for the inversion of two point-diffractor, and for the inversion of a single layer. a) The upper diffractor dominates the cross-correlogram. b) The deeper diffractor dominates the cross-correlogram. c) Inversion of the single layer.

results. The behaviour of the full waveform inversion is similar, with the exception that the shallower interface is in the right place from the first iteration because the background velocity is correct.

6.8 The Dallas model

As suggested before, the success of the full waveform and the traveltime-offset inversions depends strongly on the frequency spectrum of the wavelet and on the characteristic length of the perturbations present in the model. Therefore, a well defined synthetic model is crucial to succeed. Xu et al. (1995) reported a successful full waveform inversion of a 3D synthetic dataset. Here, I try to simulate their experiment using only a 2D image of their multilayered monocline. This is what I call the "Dallas model". Not all parameters have been described in the paper, for instance density values or preconditioning functions. If a parameter is not given, I use the one defined in my previous example.

An array of 62 receivers and 15 shots has been evenly extended along the surface. The grid consists of 129x150 points, with 10 m spatial sampling interval. The sample interval in time is 0.001 s and 1000 time steps are calculated for each trace. Gardner's formula (Gardner et al., 1974; Zhou, 1995; Zhou et al., 1995) is used to estimate the density values. The source function is the first derivative of a Gaussian function with 17 Hz as a peak frequency. The minimum and maximum velocity are restricted to 2500 m/s and 3500 m/s, respectively. The direct arrivals are muted out.

The original velocity model is described in Figure 6.36a. Once more, I started the inversion using a point-diffractor model, represented in Figure 6.36b. Figures 6.37a and 6.37b show the reconstructed image after 8 iterations of traveltime-offset and full waveform inversions, respectively. For this example, the full waveform method provides an accurate blocky result. Meanwhile, the traveltime-offset inversion resolves the interfaces but the intrabed velocities are not well recovered. Mora (1987a, 1987b) suggested the use of a "blockiness" preconditioning as a technique to enhance the low wavenumbers. The application of this preconditioning

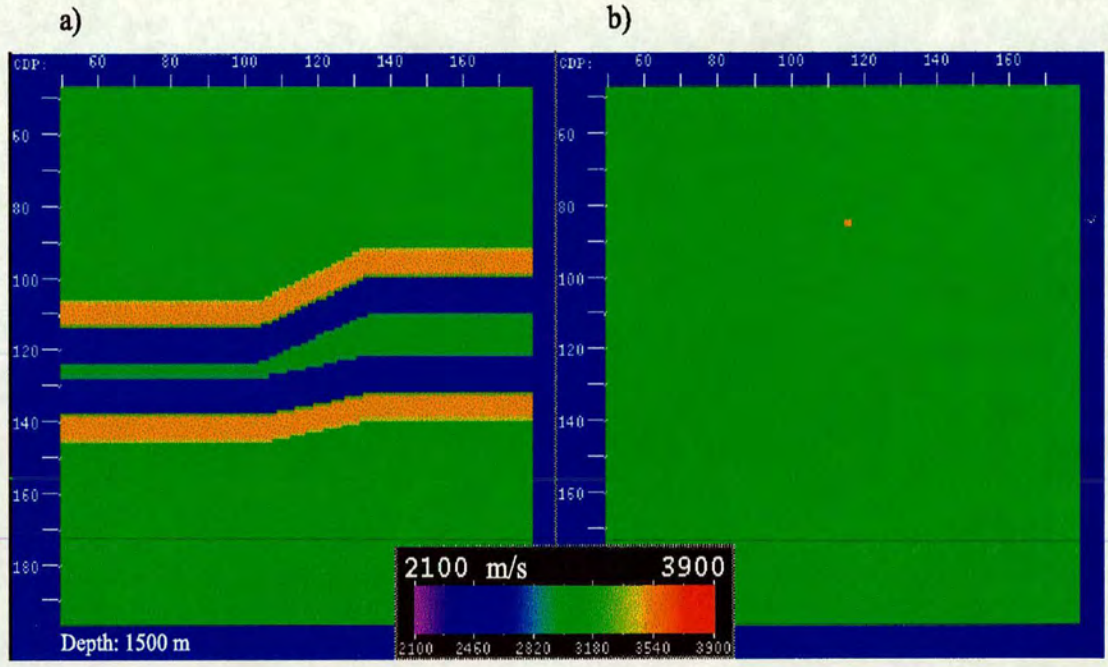


Figure 6.36: The Dallas model. a) The original synthetic velocity model. b) The starting velocity model.

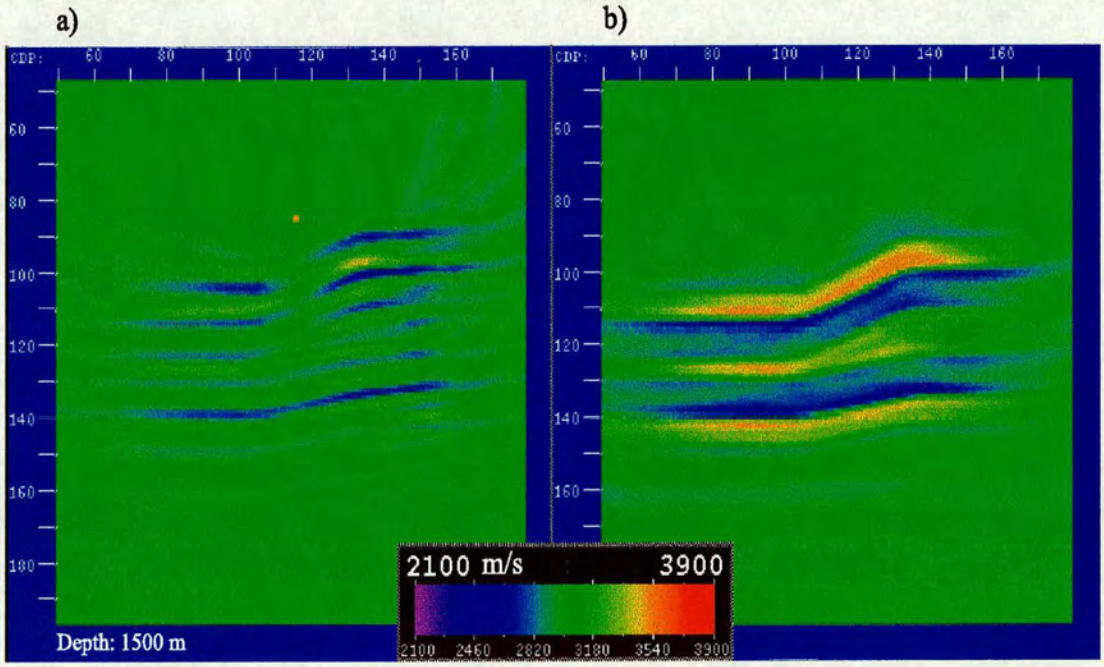


Figure 6.37: Dallas model inversions after 8 iterations. a) Traveltime-offset inversion. b) Full waveform inversion.

function may help the traveltime-offset inversion to recover the blocky aspect of the structures in the velocity model.

The output shot gathers for shot number 8 are depicted in Figure 6.38 together with a bandpass filtered version of them. It can be seen in this picture that the lowest frequencies are not present in the output shot gather from the traveltime-offset inversion. If the inversion methods differ, mainly in the residuals to be back-propagated, the origin of the differences should be in the back-propagated wavefields. Figure 6.39 depicts the fact that the traveltime residual $\delta\tau$ has a different frequency content from the actual shot gather.

It is important to note that the size of the diffractor I proposed as a starting model is 3 times smaller than the minimum wavelength in the source (62.5 m), so the wavelet hardly interacts with it. The traveltime and offset delays come from the cross-correlation between the remaining mute slopes. I tried different initial models and I obtained similar results to the one I have already shown here. As the complexity of the starting model increases, the difficulties of removing it increase as well, and a large number of extra iterations is needed. The full waveform inversion is also affected by the initial model. However, it has the advantage that the background velocity without any perturbation on it can be used as a starting model.

6.9 Spectral analysis

Pica et al. (1990) mentioned that the results obtained with gradient optimisation techniques depend strongly on the spectral content of the data. Kolb and Canadas (1986) proved the feasibility and accuracy of the least-squares inversion to solve the velocity and reflection coefficients using a frequency windowing technique. Gauthier et al. (1986) affirm that only the high frequency content of the model is recovered in a few iterations if a surface reflection dataset is used. Claerbout (1976) and Jannane et al. (1989) independently show that there is a gap in resolution between the short and long wavelengths.

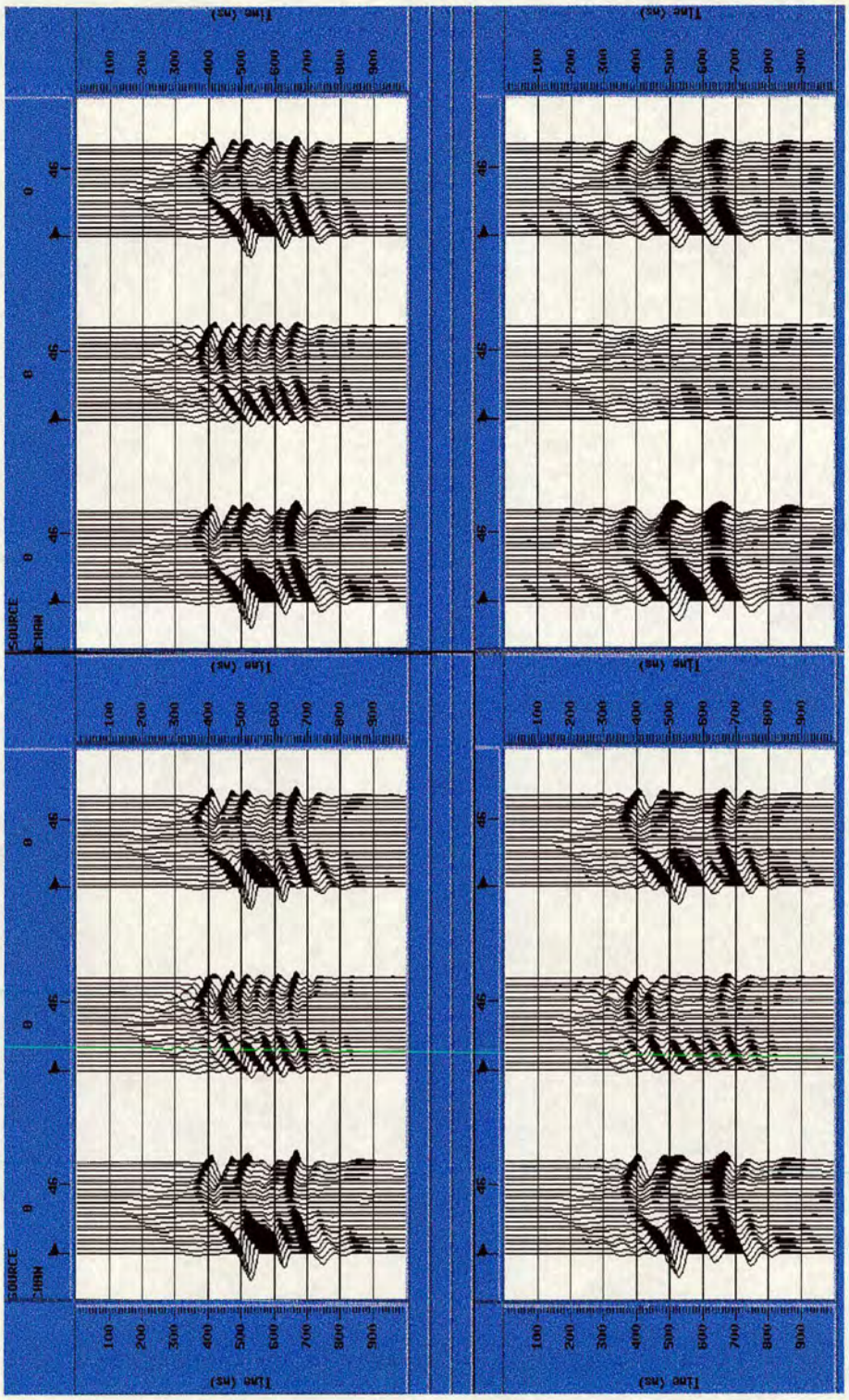


Figure 6.38: Dallas model shot gathers with a band-pass filter applied. Each graph shows the original shot gather, the traveltime-offset calculated shot gather, and the full waveform calculated shot gather, respectively. From left to right, and from top to bottom, the band-pass filters are: full frequencies, 0-1-17-20, 0-1-15-17, and 0-1-10-12, respectively.

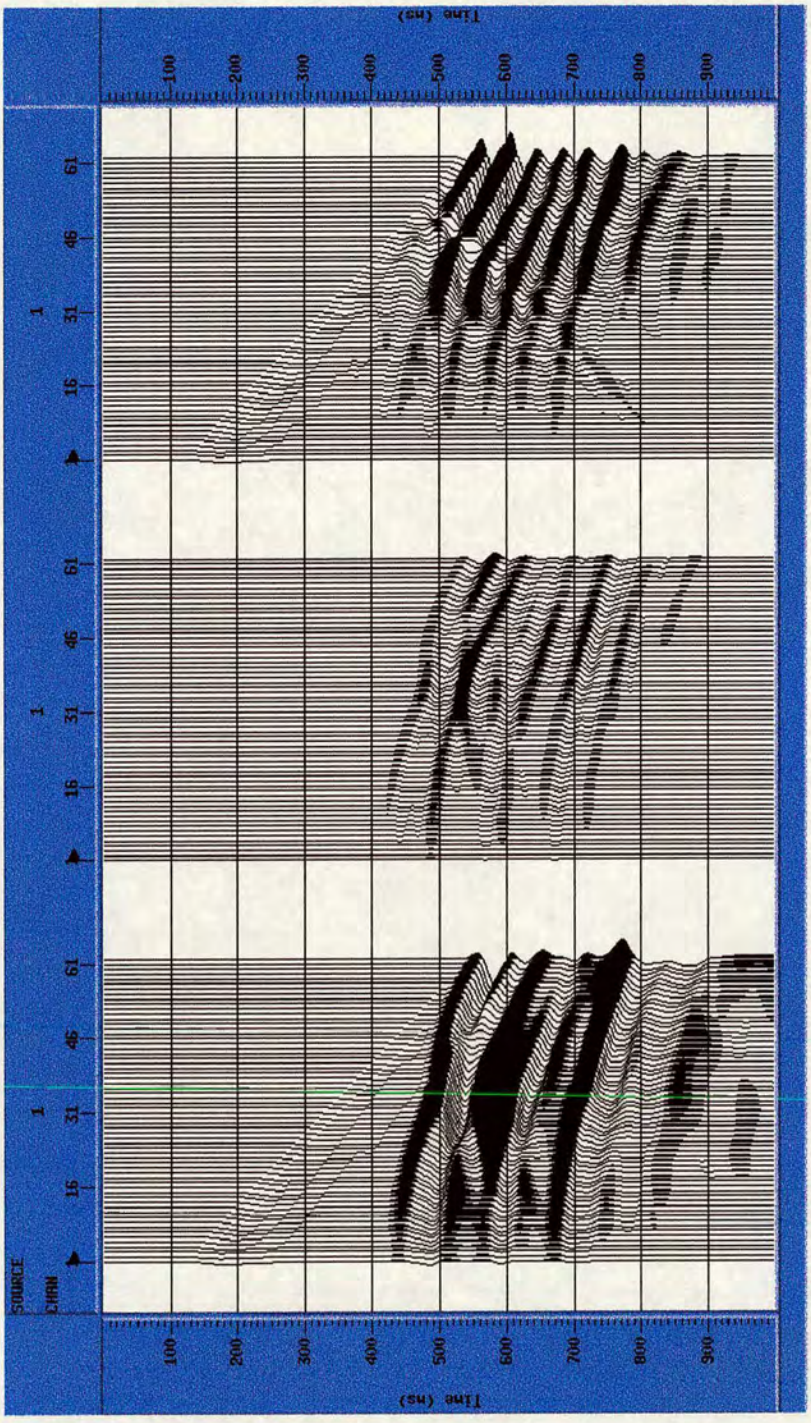


Figure 6.39: Residual shot gather. a) Original shot gather. b) Traveltime residual to be back-propagated. c) Current shot gather after 8 iterations of the traveltime-offset inversion.

Mora (1987a, 1987b, 1988, 1989) presents an extended analysis of the wavenumber spectra. He concludes that the low wavenumbers in the velocity model can be efficiently resolved when using transmission data, and that the high wavenumbers are well resolved with reflection seismic data. He shows that the inversion behaves as iterative prestack migration if reflected waves are used, and as diffraction tomography if transmitted waves are present. Using a single-frequency seismic source and assuming infinite offset ranges, he defines areas in the wavenumber spectra where the velocity perturbation can be solved. He points out that some parts of the frequency components cannot be resolved using seismic reflection datasets in real experiments due to the finite range of offsets and band-limited sources.

In my previous example, the Dallas model, I observed differences in the results from the full waveform inversion and the traveltime-offset inversion. As I attribute these differences to the frequency components of the residuals to be back-propagated, I present in this section a spectral analysis of some synthetic datasets.

The velocity model is a horizontal layer of 3500 m/s interval velocity embedded in a constant velocity medium of 3000 m/s, see Figure 6.40. I use the same parameters as in the Dallas model, including the source wavelet (first derivative of a Gaussian function with 17 Hz peak frequency). Beginning with a layer of $\lambda_\nu/4$ thickness, the spectrum for the surface-reflection configuration is shown in Figure 6.41, together with the spectra of the reflected events and the transmitted waves, respectively. For the transmission configuration I place the receivers at the bottom of the model. The frequency spectrum of the transmitted waves is similar to the spectrum of the direct+reflected arrivals, with the exception of the low frequency peak in the transmitted signals generated by the spurious signal below the event. I muted out this signal, and the new result is shown in Figure 6.42. The reflected spectrum is shifted to the right with respect to the transmitted spectrum.

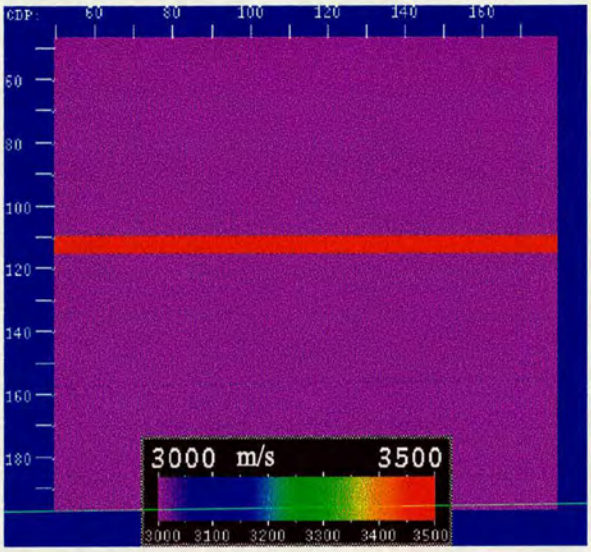


Figure 6.40: Model with a flat layer of approximately $\lambda_\nu/4$ thickness.

Figure 6.43, displays the result after increasing the thickness of the layer up to a value close to λ_ν. There are no major changes in the direct+reflected and the transmitted arrivals, but the reflected spectrum corresponds to the interference pattern between the waves reflected by both interfaces. The pattern changes as a function of the characteristic length of the model as illustrated in Figure 6.44. Therefore, the frequency spectrum of the transmitted arrivals differs from the spectrum of the reflected arrivals, and the latter depends strongly on the characteristic length of the model.

Returning to the Dallas model, Figure 6.45 shows how the mute affects the spectra. Figure 6.45a displays the direct+reflected arrivals in a single receiver. After muting out the direct arrival automatically, the remaining trace and its spectrum is shown in Figure 6.45b. A completely different result is obtained if the mute is picked manually. This lack of precision in the automatic mute affects mainly the near-offset traces. However, as these particular traces have low amplitudes, the spectrum of the complete shot gather is not altered (Figure 6.46).

The traveltime-offset and the full waveform inversions of the Dallas model are

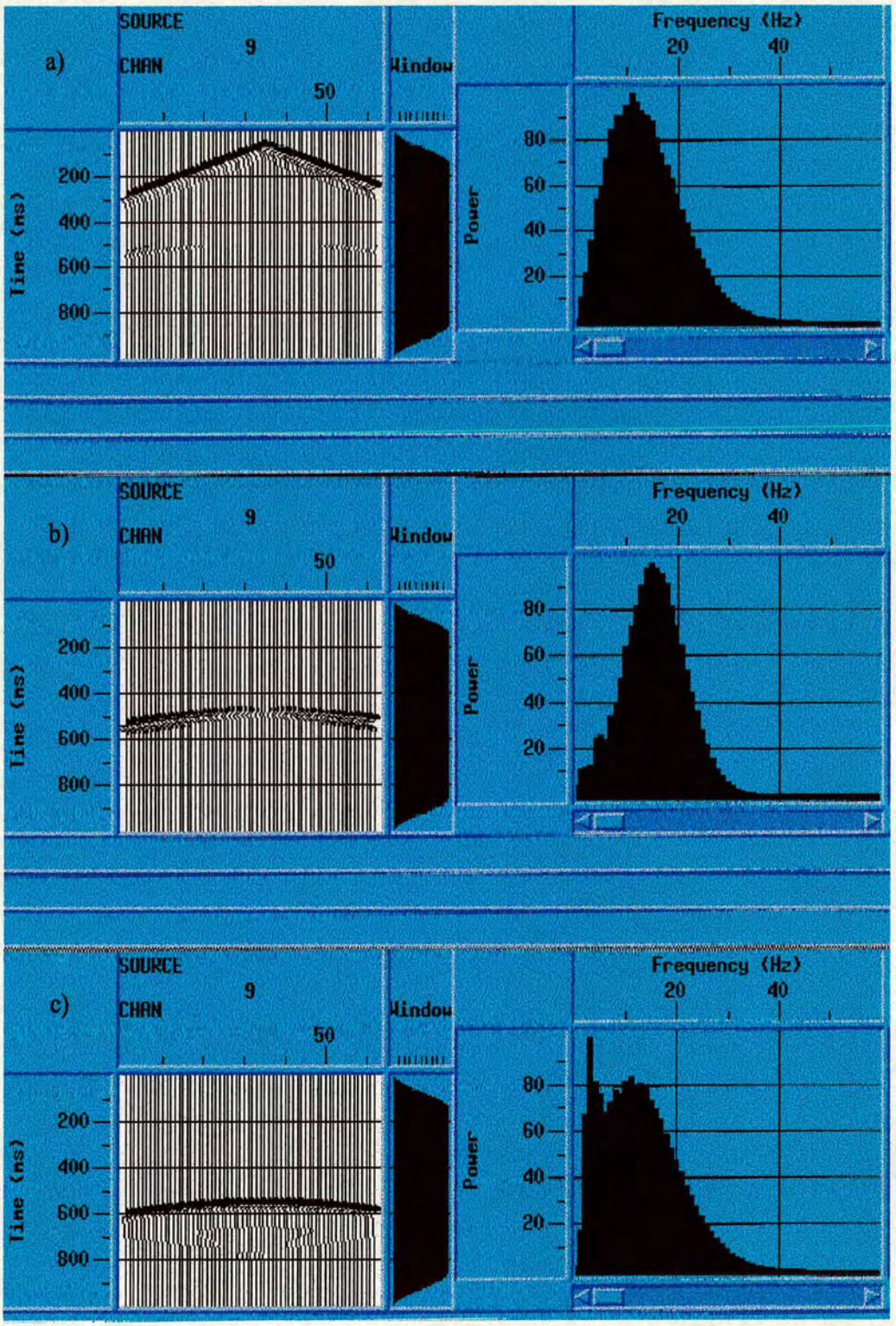


Figure 6.41: Spectra. a) Direct+reflected arrivals. b) Reflected arrivals. c) Transmitted arrivals.

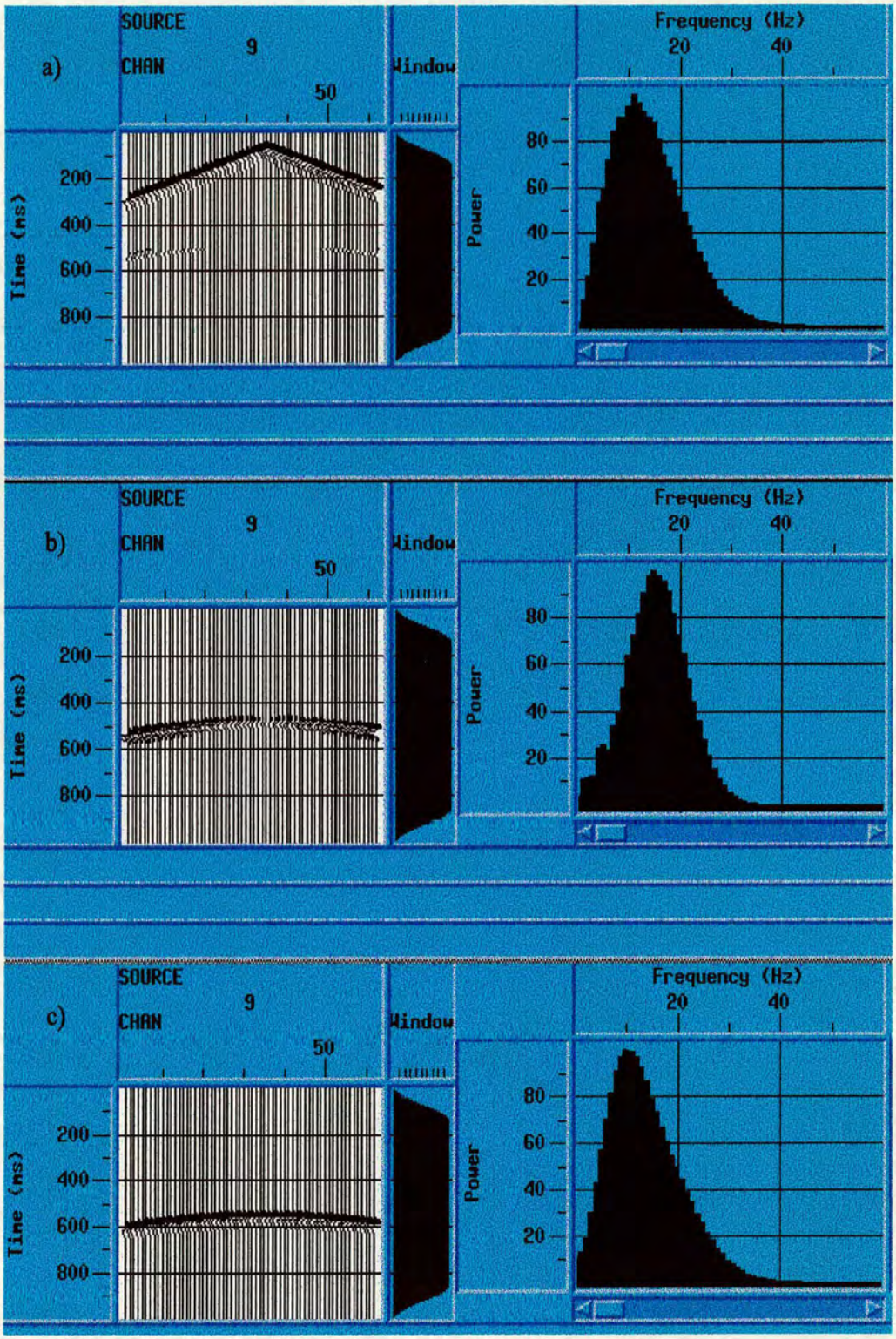


Figure 6.42: Same as figure 6.40, but with the spurious signals below the transmitted arrivals muted out.

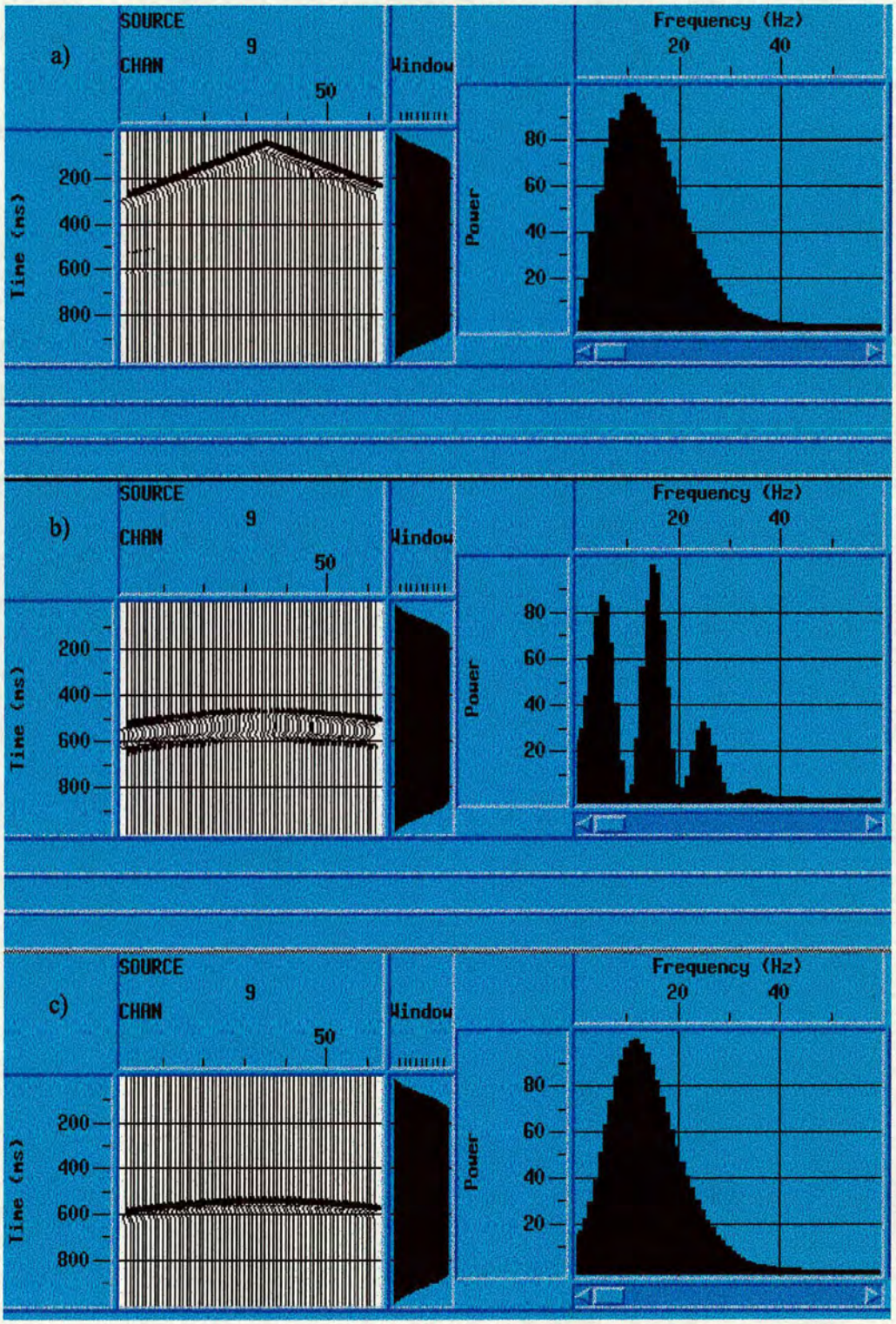


Figure 6.43: Spectra for the model with a flat layer of, approximately, λ_ν thickness. a) Direct+reflected arrivals. b) Reflected arrivals. c) Transmitted arrivals.

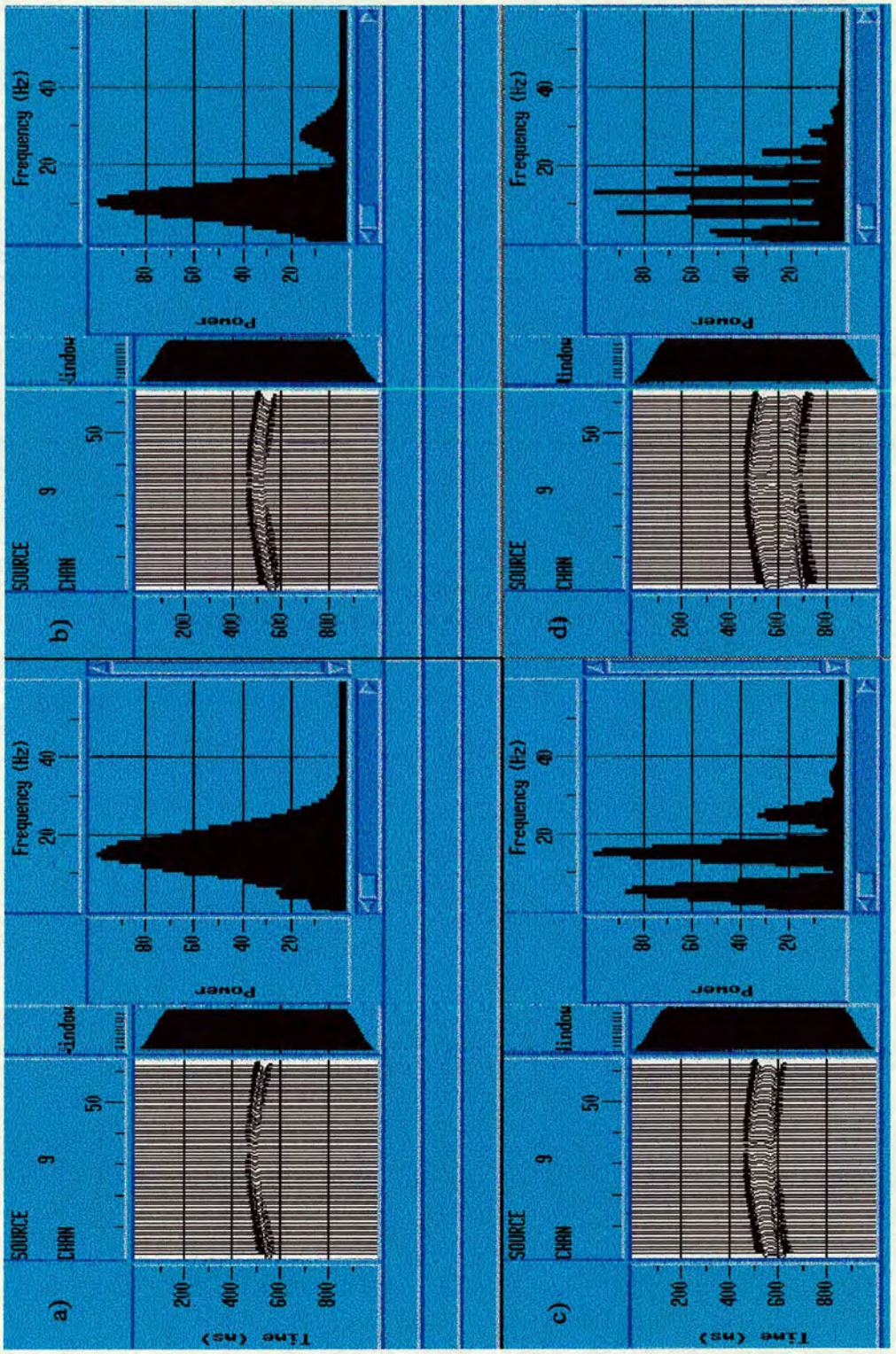


Figure 6.44: Spectra as a function of the thickness of the layer. a) $\lambda_\nu/4$. b) $\lambda_\nu/2$. c) λ_ν. d) $2\lambda_\nu$.

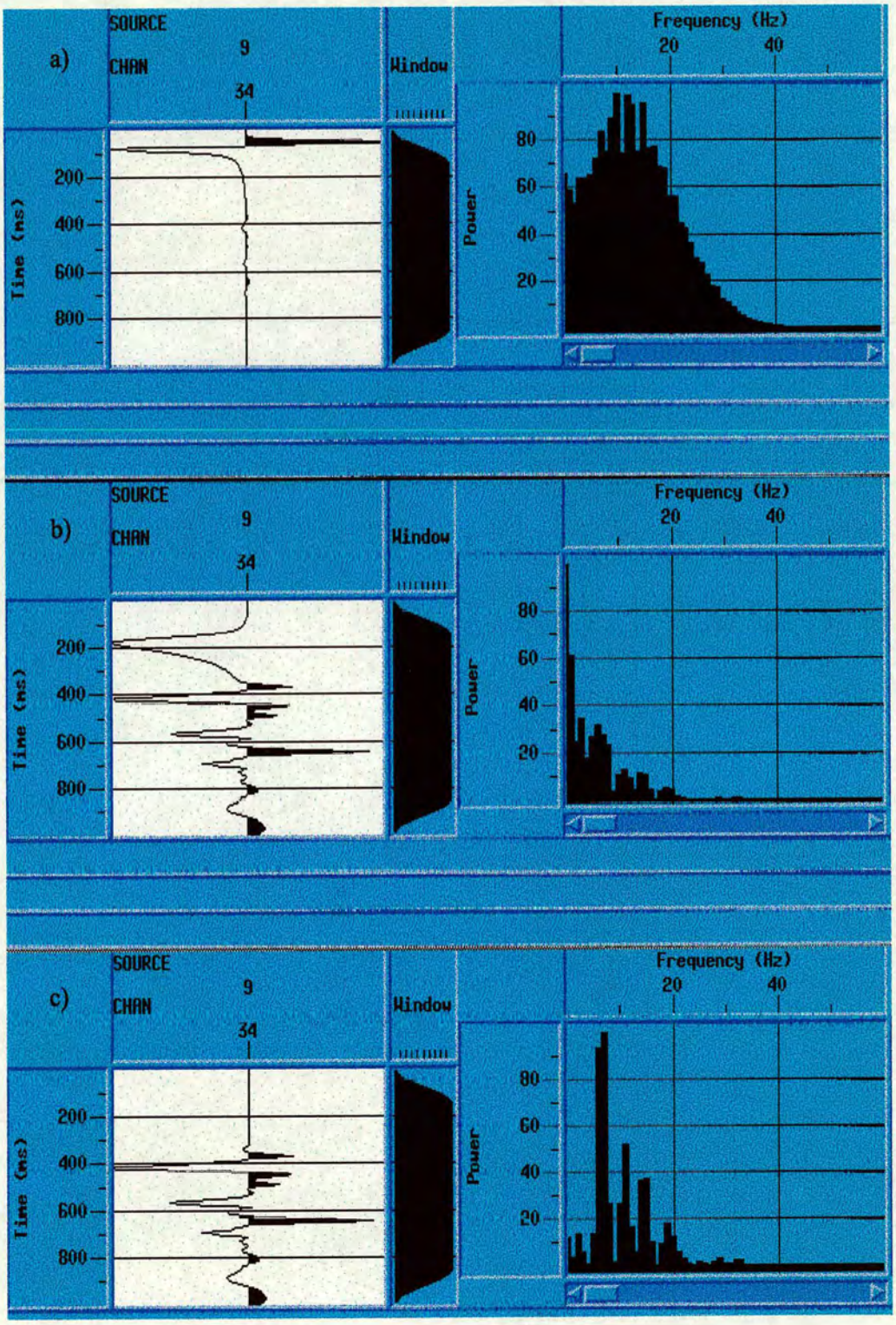


Figure 6.45: Spectrum of one trace in the Dallas model. a) Observed trace. b) After automatic muting. c) After manual muting.

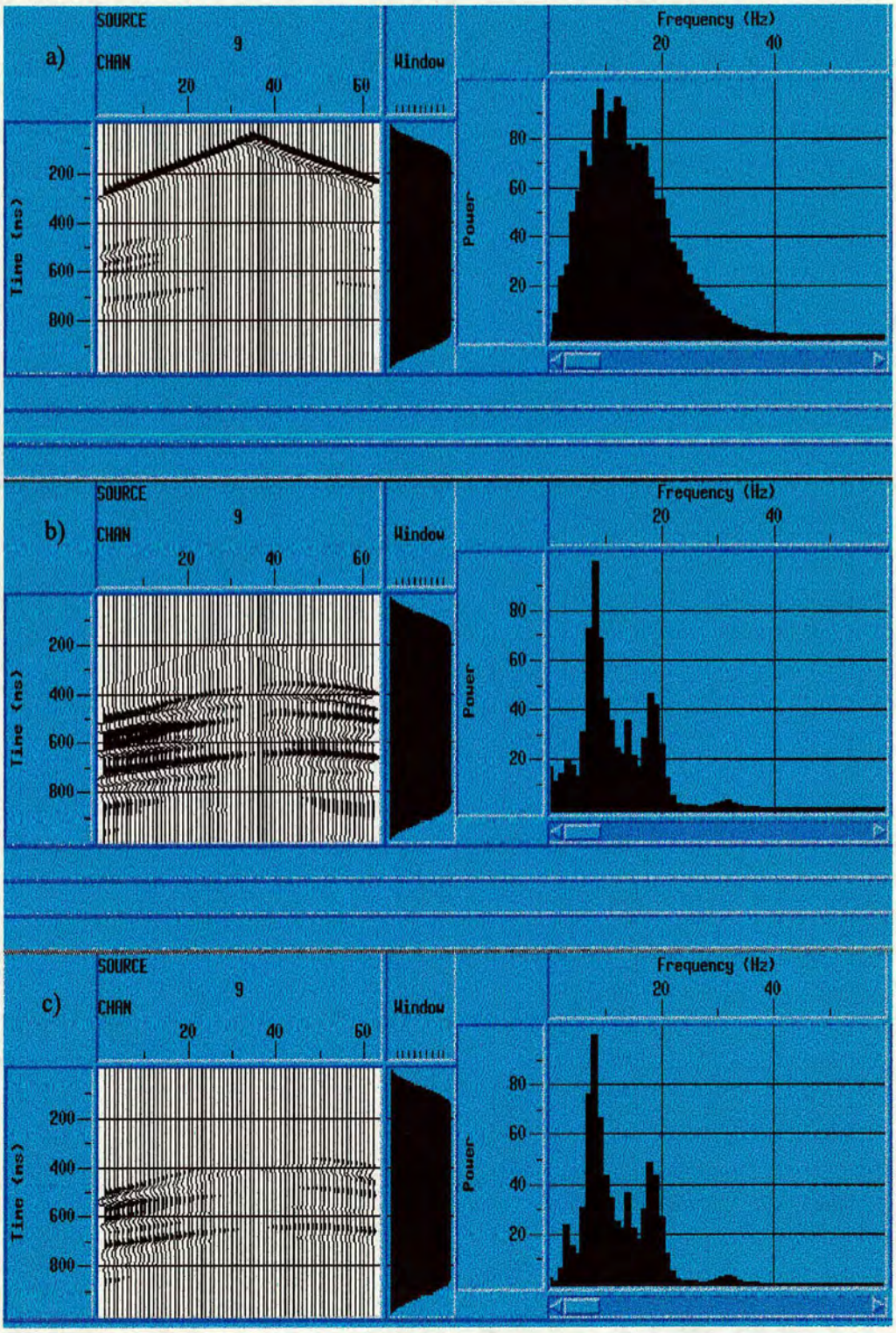


Figure 6.46: Spectrum of one shot in the Dallas model. a) Observed shot gather. b) After automatic muting. c) After manual muting.

carried out using reflected datasets. The spectrum of one of the observed shot gathers is shown in Figure 6.47. The inputs are the same for both inversions, but the frequency content of the results is different. The difference is generated by the spectra of the residuals to be back-propagated. In Figure 6.47, I display the spectrum of the traveltime residual $\delta\tau$ together with the spectrum of the seismogram residuals δp. The offset residual $\delta\sigma$ is depicted in Figure 6.48.

The first derivative of the observed shot gather with respect to time or offset, increases the high frequency content of the signals. This is why the calculated shot gather after 8 iterations of traveltime-offset inversion has higher frequencies on it than the observed shot gather (Figure 6.49). Therefore, the traveltime-offset inversion correctly places the dominant event, recovers the velocity above it, and reconstructs the interfaces of the remaining events. Nevertheless, the blocky variations in the velocity model cannot be efficiently resolved. In contrast, the shot gather residual and the observed shot gather have similar spectra (assuming the initial velocity model is the original background velocity). The middle frequencies in this spectrum produce the blocky effect observed in the velocity model after applying 8 iterations of full waveform inversion (see Figure 6.37).

6.10 Inverting for the background velocity

As mentioned by Luo and Schuster (1991), Xu et al. (1995), Pica et al. (1990), and Mora (1987a, 1987b), among others, a fairly accurate initial velocity model containing the low wavenumbers is required in gradient-based methods. This means that the starting velocity model should reproduce the low frequency components of the actual velocity field, and the structures in the real model are considered as perturbations to be introduced in this initial model. I now consider what happens when these perturbations are included in the initial guess but the background velocity is wrong.

Figure 6.50b outlines a starting model that differs from the real one only in the velocity of the first layer: the original velocity is 3000 m/s and the initial one is 2800 m/s. After 5 iterations, the estimated velocity in the upper layer is

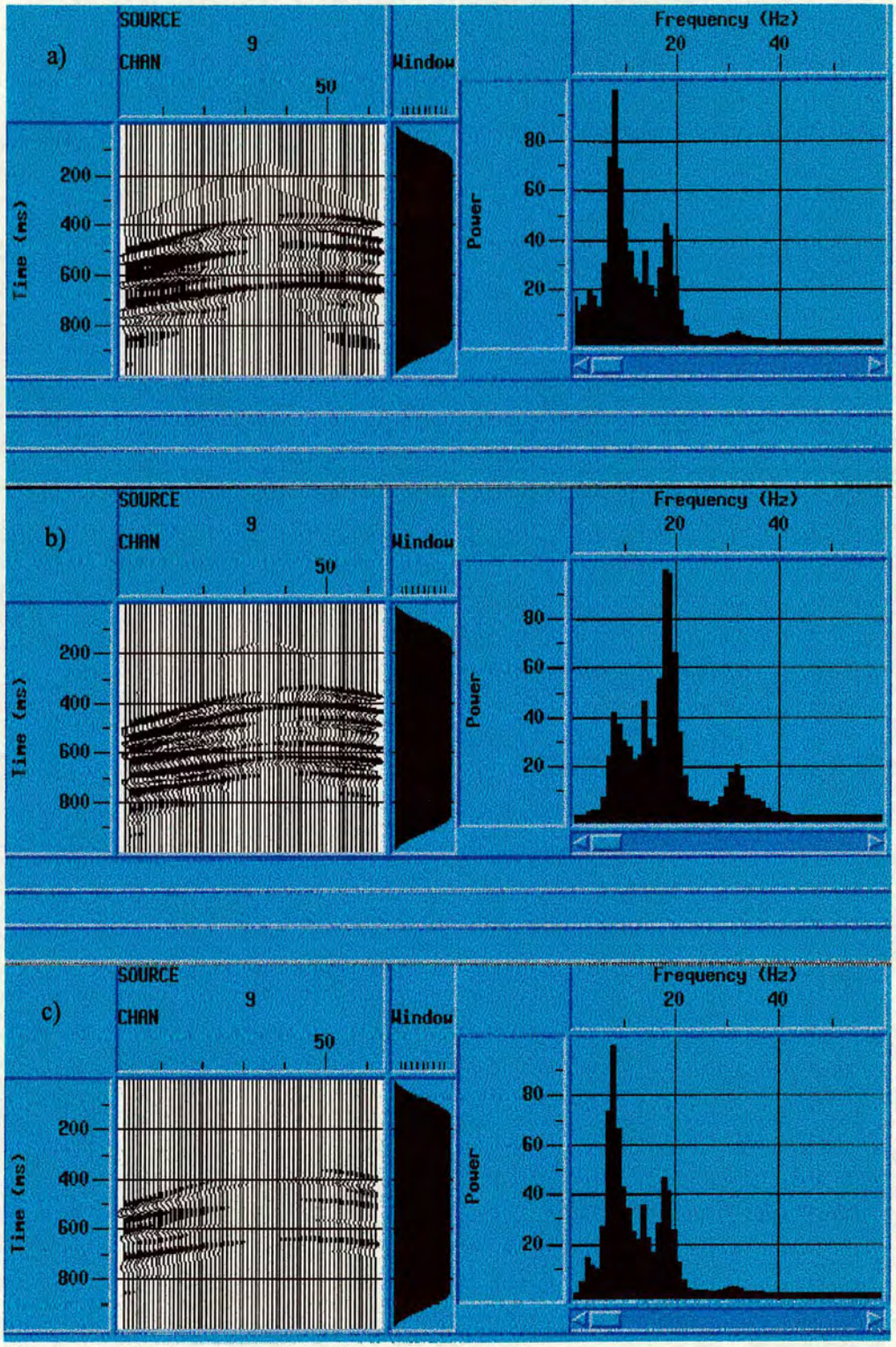


Figure 6.47: Spectra of the traveltime residual. a) Original shot gather. b) Traveltime residual shot gather ($\delta\tau$) c) Seismogram residual (δp).

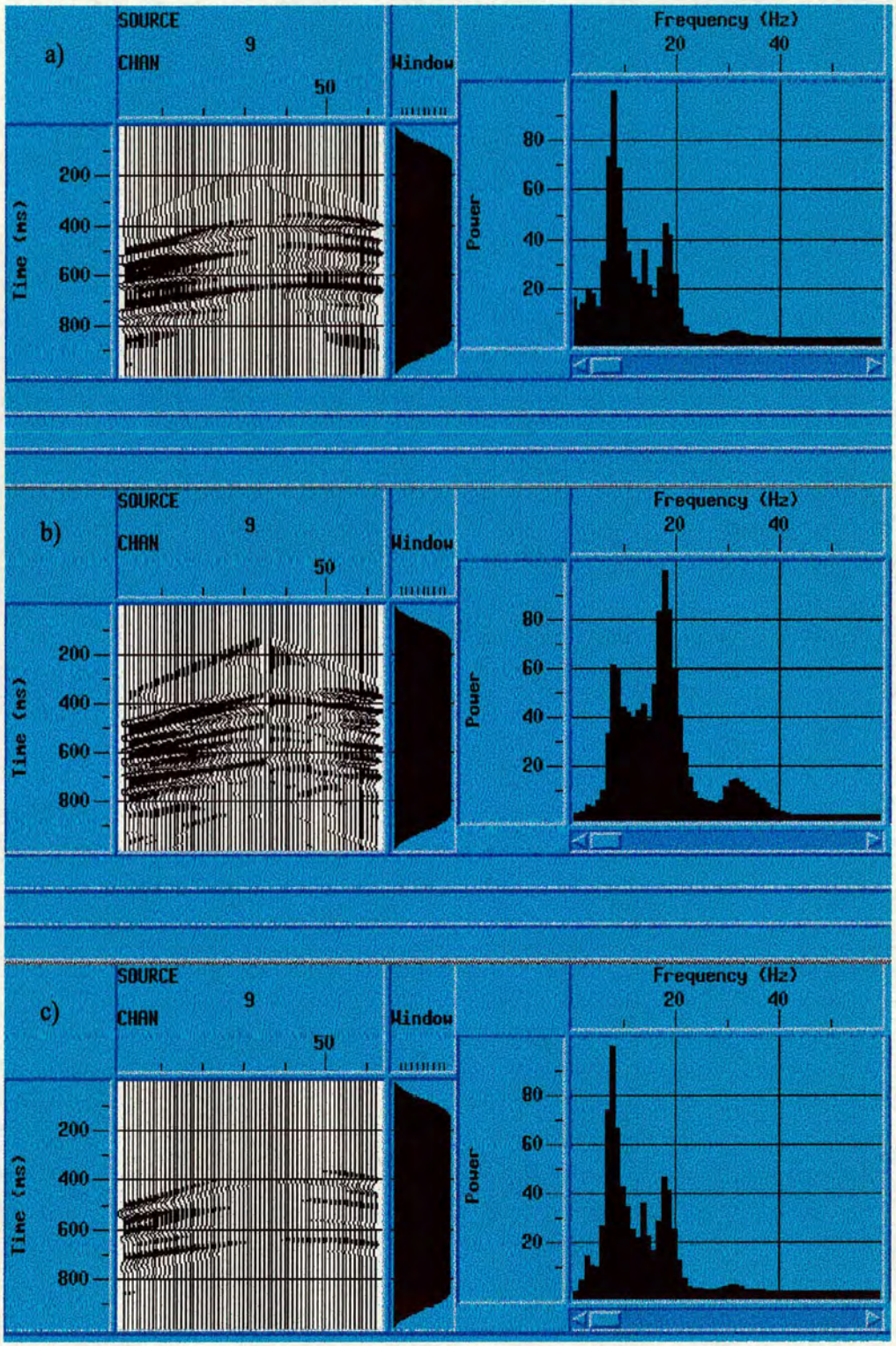


Figure 6.48: Spectrum of offset residual. a) Original shot gather. b) Offset residual shot gather ($\delta\sigma$) c) Seismogram residual (δp).

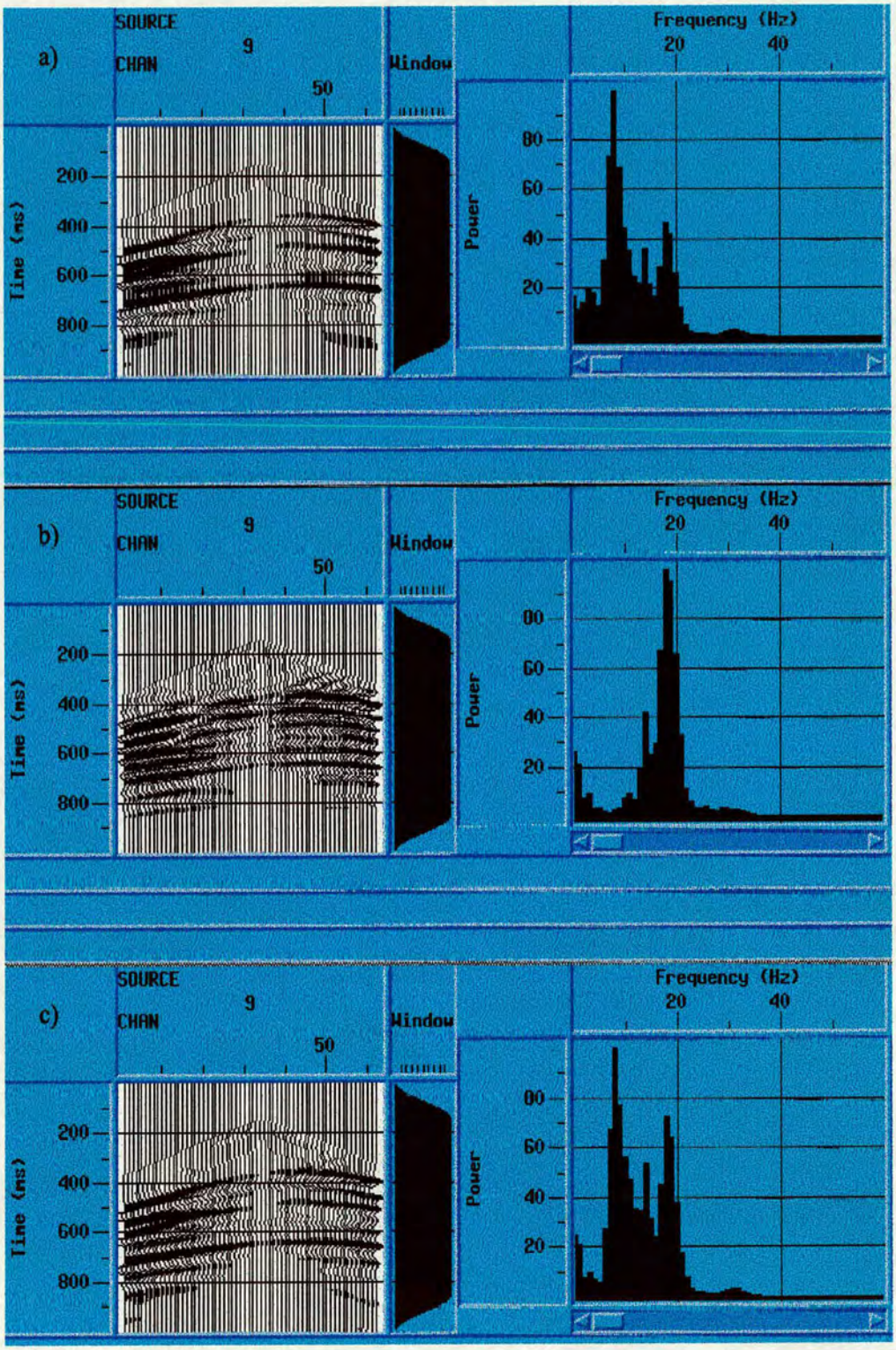


Figure 6.49: Spectra after inversions. a) Original shot gather. b) Traveltime-offset inversion after 8 iterations. c) Full waveform inversion after 8 iterations.

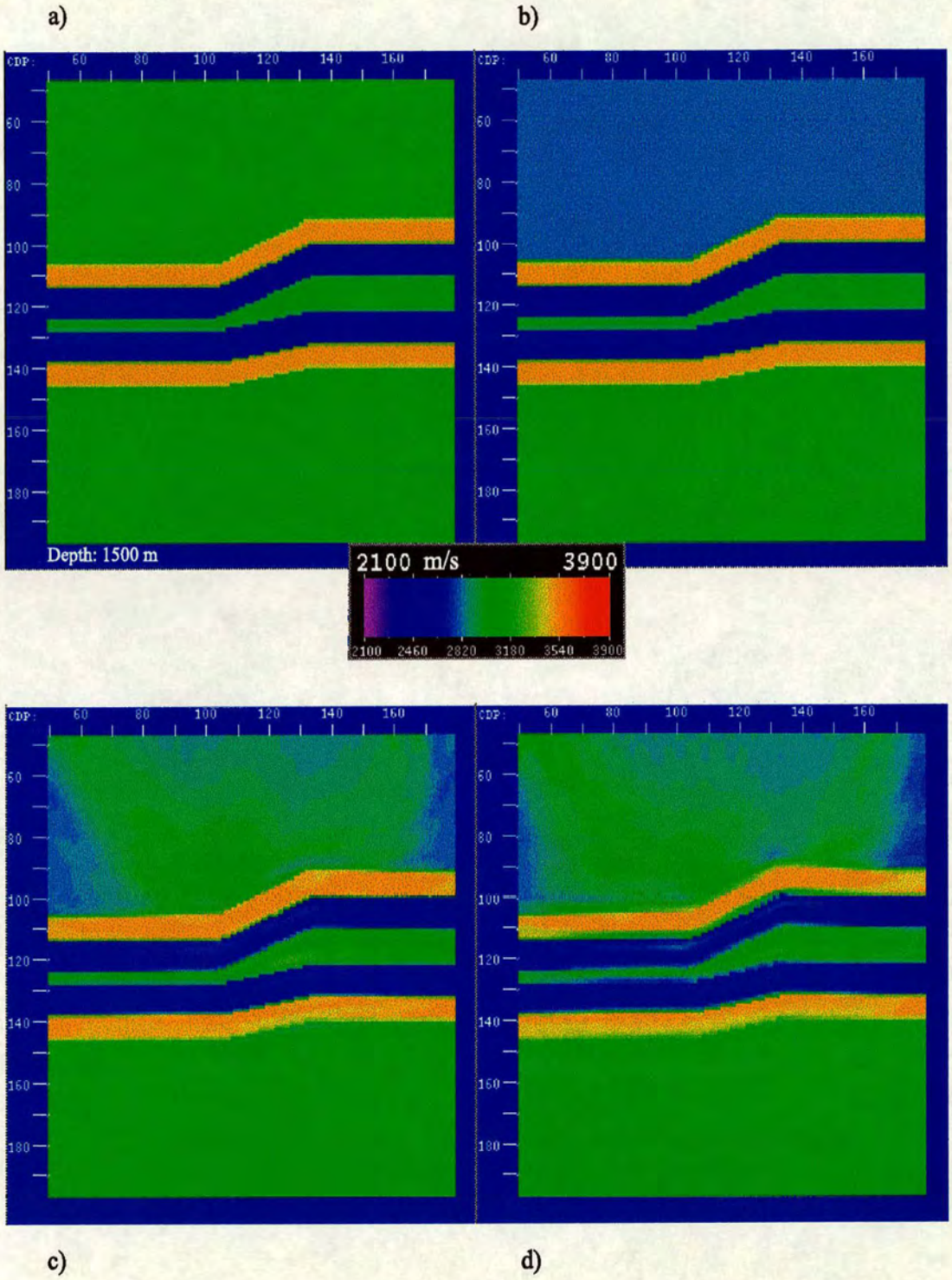


Figure 6.50: Inversion using 2800 m/s as initial velocity in the upper layer. a) Original velocity model. b) Initial velocity model. c) Traveltime-offset inversion after 5 iterations. d) Full waveform inversion after 5 iteration s.

almost correct but there is a slight movement of the reflectors to the upper part to compensate the traveltimes. This effect is more obvious in the full waveform than in the traveltime-offset inversion.

The full waveform inversion reaches a critical point if the starting model is moderately far from the actual velocity model. In Figure 6.51, I display the result of using 2600 m/s in the upper layer as a starting model. The full waveform inversion converges to a local minimum. Meanwhile, after 10 iterations the traveltime-offset inversion has recovered part of the original velocity field and keeps the structural features almost in the right position.

6.11 Summary

The new traveltime-offset inversion combined with the full waveform inversion provides an accurate reconstruction of the velocity model when using transmitted arrivals from crosshole datasets. For surface reflection geometries, the traveltime-offset inversion reconstructs the main event and the velocity values above it (using the traveltime and offset information). The remaining interfaces are recovered as well. However, the proposed method cannot reproduce, in a few iterations, the low wavenumber associated with the blocky velocity effect. This is because the derivatives involved in the pseudo-residuals to be back-propagated move the spectrum towards higher frequencies.

The frequency windowing technique helps to avoid getting stuck in a local minimum. Nevertheless, as the traveltime and offset delays diminish rapidly during the inversion of the first frequency band, the procedure shows a lack of efficiency when trying to remove the initial model.

By taking into account the moveout between traces, the convergence of the traveltime-offset inversion method is more stable than the convergence of the standard traveltime method. The latter method can diverge if small changes are made in the step length values. Also, by keeping the moveout of the original shot gathers, the shapes of the interfaces are well defined during the traveltime-offset

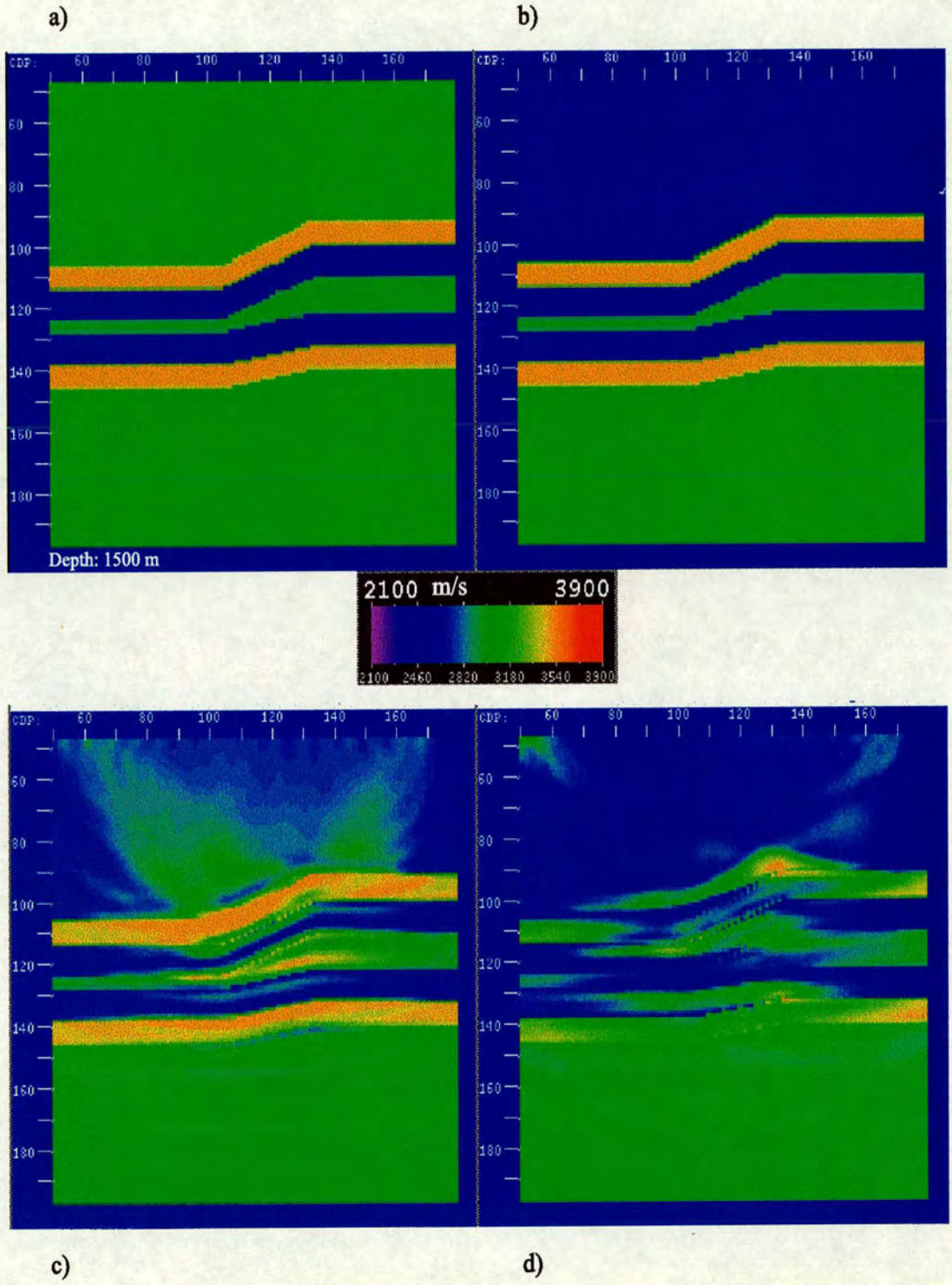


Figure 6.51: Inversion using 2600 m/s as initial velocity in the upper layer. a) Original velocity model. b) Initial velocity model. c) Traveltime-offset inversion after 10 iterations. d) Full waveform inversion after 10 iterations.

inversion. Using synthetic datasets, I have shown the feasibility of the proposed method and how certain parameters modify the results.

Chapter 7

Conclusions

In this chapter I summarise the main results of this dissertation, and suggest ideas to improve the results of the new traveltime-offset inversion method along with possible directions of future works.

In order to reconstruct the compressional wave velocity field, three different wave equation inversion methods have been described and implemented on a Cray T3D computer. These consist of: the standard traveltime inversion, the full waveform inversion, and the new traveltime-offset inversion. All of them are based on the acoustic wave-equation, and they update the current velocity model by correlating the corresponding back-propagated pseudo-residuals with the forward wavefield. A finite-difference scheme is used to propagate the wavefields.

The steepest descent direction of the misfit function is used to update the velocities and, in practice, a preconditioning function is applied to accelerate the convergence. These methods are free of approximations and traveltime picking is unnecessary. In theory, all waveforms can be considered, but the resolution is conditioned by the frequency content of the data to be inverted, and by the starting velocity model.

The standard traveltime inversion successfully reconstructs the low frequency components of a faulted velocity model using transmitted arrivals from a crosshole dataset. This method fails, however, when reflected or diffracted arrivals from surface reflection datasets are considered. Artifacts are present in the updated velocity model, and the spikes observed in the traveltime residuals are sensitive to the step length values even when the residuals are small. The convergence is

unstable, and as the traces are shifted independently in a common shot gather the moveout between traces is not taken into account during the inversion.

The full waveform inversion recovers the high-wavenumber components of the velocity field, but a starting velocity model containing the low-frequency content is required to guarantee the convergence of the method to the global minimum. Combined with the standard traveltime inversion, it succeeds in the reconstruction of a synthetic crosshole dataset. The artifacts I described when inverting surface reflection datasets are related to the frequency spectrum of the wavelet and to the characteristic wavelength of the features present in the velocity model. The convergence of this method is stable and the residuals oscillate if the step length value is not appropriate.

I propose a new wave-equation inversion method (traveltime-offset inversion) that includes the offset dependence of the traces in a shot gather with reflected or diffracted arrivals. This is done by introducing a bi-dimensional (offset-time) cross-correlation function as a connective function. The misfit function takes into account the offset shifts as well as the traveltime delays. Two different pseudo-residuals (offset and traveltime) are generated, and the velocity model is updated by correlating the forward wavefield with the fields computed by reverse time propagation of these pseudo-residuals acting as sources at the receiver locations.

As all traces in an observed shot gather are rigidly moved, the moveout is maintained. This fact allows a rapid reconstruction of the shape of the reflectors and, therefore, there is an improvement in the rate of convergence. This also explains why the convergence of this method is more stable than with the standard traveltime inversion.

The new method combined with the full waveform inversion reproduces the faulted synthetic model with crosshole configuration. It also succeeds in inverting surface reflection datasets while the standard traveltime inversion fails. The main event and the velocity above it are recovered when datasets with surface geometry are used. The remaining interfaces are well defined, but not correctly placed

in a few iterations. The blocky effect associated with the low wavenumbers is not reproduced because the derivatives involved in the pseudo-residuals move the spectrum towards higher wavenumbers.

The full waveform inversion fails when the background velocity in the starting model differs by more than 10% from the original one, while the new traveltime-offset inversion recovers the original velocity field in the upper part of the model using the traveltime and the offset information. Thus, a good strategy continues to be the reconstruction of the ultra-low frequency content of the model and the definition of the interfaces using the traveltime-offset inversion, followed by the full waveform inversion to incorporate the blocky effect in the interval velocities.

However, the new traveltime-offset inversion requires some refinement. For instance, a progressive downward inversion combined with the frequency windowing technique as suggested by Kolb et al. (1986) can be included in the new method. Different preconditioning functions, including the "blockiness" preconditioning suggested by Mora (1987a, 1987b), should be tested in order to improve the results as well as to speed up the convergence. The algorithm may not require the full waveform inversion after incorporating the blockiness preconditioning in the traveltime-offset inversion, and the ambiguity velocity-depth may be solved. Mora (1987b) also suggests that the simultaneous convergence of all wavenumbers can be achieved in his inversion by "boosting" the low wavenumbers. Then, a "boosting" of specific frequency band should be considered here.

A more precise calculation of the step length value is needed to improve the traveltime-offset inversion. Tarantola (1986, 1987), Mora (1987a, 1988), and Pica et al. (1990) suggest that a search for this value must be done for highly non-linear misfit functions. Another parameter to be considered is a weighting factor that balances out the contribution of the offset and the time residuals. For instance, the reduction of the traveltime delays may be accelerated by increasing the contribution of the pseudo-traveltime residual to the gradient.

After improving the new traveltime-offset inversion method the following step

would be the application of the new method to the inversion of 2D real datasets. A logical extension would be to use the elastic wave-equation and the elastodynamic finite-difference modelling to reconstruct P- and S-wave velocities. Another interesting topic to be considered in the future would be the extension of the method to process 3D datasets. In this case a tri-dimensional cross-correlation function should be used as the connective function.

Bibliography

Aki, K., and Richards, P., 1980, *Quantitative seismology*, Freeman Press.

Berkhout, A. J., 1985, *Imaging of acoustic energy by wave field extrapolation: theoretical aspects*, Elsevier Science Publishers.

Biondi, B., 1992, *Velocity estimation by beam stack*, Geophysics, **57**, 1034-1047.

Bishop, T., Bube, K., Cutler, R., Langan, R., Love, P., Resnick, J., Shuey, R., Spindler, D., and Wyld, H., 1985, *Tomographic determination of velocity and depth in laterally varying media*, Geophysics, **50**, 903-923.

Bracewell, R. N., 1986, *The Fourier transform and its applications*, McGraw-Hill Book Company.

Brook, D., and Wynne, R. J., 1988, *Signal processing: principles and applications*, Edward Arnold.

Chavent, G., and Jacewitz, C., A., 1995, *Determination of background velocities by multiple migration fitting*, Geophysics, **60**, 476-490.

Claerbout, J. F., 1976, *Fundamentals of geophysical data processing*, McGraw-Hill Book Co.

Dines, K., and Lytle, R., 1979, *Computized geophysical tomography*, Proc. IEEE, **67**, 1065-1072.

Gardner, G. H. F., Gardner, L. W., and Gregory, A. R., 1974, *Formation velocity and density - The diagnostic basis of stratigraphic traps*, Geophysics, **39**, 770-780.

Gauthier, O., Virieux, J., and Tarantola, A., 1986, *Two-dimensional nonlinear inversion of seismic waveforms: numerical results*, Geophysics, **51**, 1387-1403.

Gersztenkorn, A., and Scales, J., 1988, *Smoothing seismic tomograms with alpha-trimmed means*, Geophysical Journal, **92**, 67-72.

Gonzalez, R. C., and Wintz, P., 1977, *Digital image processing*, Addison-Wesley Publishing Company.

Ivansson, S., 1985, *A study of methods for tomographic velocity estimation in the presence of low-velocity zones*, Geophysics, **50**, 969-988.

Jannane, M., Beydoun, W., Crase, E., Cao, D., Koren, Z., Landa, E., Mendes, M., Pica, A., Noble, M., Roeth, G., Singh, S., Snieder, R., Tarantola, A., Trezeguet, D., and Xie, M., 1989, *Wavelengths of earth structures that can be resolved from seismic reflection data*, Geophysics, **54**, 906-910.

Jeannot, J. P., and Berranger, I., 1994, *Ray-mapped focusing: a migration velocity analysis for Kirchhoff prestack depth imaging*, 64th SEG meeting, Expanded Abstracts, 1326-1329.

Jones, I. F., 1994 *3D velocity-depth model building-layer stripping versus map-migration*, 56th EAEG meeting, Expanded Abstracts, B044.

Kinsler, L. E., Frey, A. R., Coppens A. B., and Sanders, J. V., 1982, *Fundamentals of acoustics*, John Wiley & Sons.

Kolb, P., and Canadas, G., 1986, *Least squares inversion of pre-stack data: simultaneous identification of density and velocity*, 56th SEG meeting, Houston, Expanded Abstracts, 604-607.

Kolb, P., Collino, F., and Lailly, P., 1986, *Pre-stack inversion of a 1-D medium*, Proceedings of the IEEE, **74**, 498-508.

Landa, E., Thore, P., Sorin, V., and Koren, Z., 1991, *Interpretation of velocity estimates from coherency inversion*, Geophysics, **56**, 1377-1383.

Leggett, M., Goulty, N. R., and Kragh, J. E., 1993, *Study of traveltime and amplitude time-lapse tomography using physical model data*, Geophysical Prospecting, **41**, 599-619.

Liao, Q., and McMechan, G. A., 1996, *Multifrequency viscoacoustic modeling and inversion*, Geophysics, **61**, 1371-1378.

Liu, Z., and Bleistein, N., 1995 *Migration velocity analysis: theory and a iterative algorithm*, Geophysics, **60**, 142-153.

Luo, Y., and Schuster, G., 1991, *Wave-equation traveltime inversion*, Geophysics, **56**, 645-653.

MacKay, S., and Abma, R., 1992, *Imaging and velocity estimation with depth-focusing analysis*, Geophysics, **57**, 1608-1622.

Marion, A., 1991, *An introduction to image processing*, Chapman and Hall.

Message Passing Interface Forum, 1994, *MPI: A message-passing interface standard*, University of Tennessee.

Mora, P., 1987a, *Nonlinear two-dimensional elastic inversion of multioffset seismic data*, Geophysics, **52**, 1211-1228.

Mora, P., 1987b, *Elastic wavefield inversion for low and high wavenumbers of the P- and S-wave velocities, a possible solution*, in Deconvolution and inversion, Wortington, M., Blackwell Scientific Publications.

Mora, P., 1988, *Elastic wave-field inversion of reflection and transmission data*, Geophysics, **53**, 750-759.

Mora, P., 1989, *Inversion = migration + tomography*, Geophysics, **54**, 1575-1586.

Morse, P. M., and Feshbach, H., 1953, *Methods of theoretical physics*, McGraw-Hill Book Company.

Pratt R. G., and Goulty, N. R., 1991, *Combining wave-equation imaging with traveltime tomography to form high-resolution images from crosshole data*, Geophysics, **56**, 208-224.

Pica, A., Diet, J. P., and Tarantola, A., 1990, *Nonlinear inversion of seismic reflection data in a laterally invariant medium*, Geophysics, **55**, 284-292.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 1994, *Numerical recipes in fortran*, Cambridge University Press.

Ristow D., and Rühl T., 1994, *Fourier finite-difference migration*, Geophysics, **59**, 1882-1893.

Rowbotham P. S., and Goulty, N. R., 1993, *Imaging capability of cross-hole seismic reflection surveys*, Geophysical Prospecting, **41**, 927-941.

Savic, Milos., 1995, *Ultrasonic scattering from a hydraulic fracture: theory, computation and experiment*, Ph.D. thesis, Delft University of Technology.

Sheriff, R. E., and Geldart, L. P., 1995, *Exploration seismology.* Cambridge University Press.

Stoffa, P. L., Fokkema, J. T., de Luna Freires, R. M., and Kessinger, W. P., 1990, *Split-step Fourier migration*, Geophysics, **55**, 410-421.

Stork, C., and Clayton, R., 1986, *Analysis of resolution between ambiguous velocity and reflector position for traveltime tomography*, 56th SEG meeting, Houston, Expanded Abstracts, 545-550.

Sun, R., and McMechan G. A., 1992, *2-D full-wavefield inversion for wide-aperture, elastic, seismic data*, Geophys. J. Int., **111**, 1-10.

Symes, W. W., and Carazzone, J. J., 1991, *Velocity inversion by differential semblance optimization*, Geophysics, **56**, 654-663.

Tanis, M. C., and Stoffa, P. L., 1997, *Parallel implementation of 3-D split-step Fourier depth migration algorithm on T3E*, 67th SEG meeting, Dallas, Expanded Abstracts, 1433-1436.

Tarantola, A., 1984, *Inversion of seismic reflection data in the acoustic approximation*, Geophysics, **49**, 1259-1266.

Tarantola, A., 1986, *A strategy for nonlinear elastic inversion of seismic reflection data*, Geophysics, **51**, 1893-1903.

Tarantola, A., 1987, *Inverse problem theory.* Elsevier Sci. Publ.

Varela, C. L., Stoffa, P. L., and Sen, M., 1994, *Migration misfit and reflection tomography: criteria for prestack migration velocity estimation in laterally varying media*, 64th SEG meeting, Expanded Abstracts, 1347-1350.

Versteeg, R., J., and Marfurt, K. J., 1994, *Exploring model space on the cheap: a hybrid approach to velocity model estimation*, 64th SEG meeting, Expanded Abstracts, 1202-1204.

Xu, T., McMechan, G. A., and Sun, R., 1995, *3-D prestack full-wavefield inversion*, Geophysics, **60**, 1805-1818.

Zhou, Ch., 1995, *Acoustic and elastic traveltime and waveform inversion of cross-well seismic data*, Ph.D. thesis, The University of Utah.

Zhou, Ch., Cai, W., Luo, Y., Schuster, G., and Hassanzadeh, S., 1995, *Acoustic wave-equation traveltime and waveform inversion of crosshole seismic data*, Geophysics, **60**, 765-773.

Appendix A

Split-step Fourier and Fourier finite-difference migrations

A.1 Motivation

In a large seismic survey only a few lines are considered for velocity inversion in order to keep costs low. Consequently, an accurate migration method is needed to complete and correct the image of the structure beneath the surface. The automated velocity analysis scheme, shown in Figure 1.1, includes a migration algorithm together with a migration misfit criterion. In that scheme, the calculated seismogram obtained by forward modelling of the initial velocity model is compared with the real data. If the misfit criterion for the velocity analysis is not satisfied, the velocity model is perturbed until the minimum error is reached.

The migration algorithm is applied using the last updated velocity model. If the image fulfils the expectations of the interpreters the procedure can be stopped; otherwise a different misfit criterion has to be satisfied and the velocity field is updated again. Two different paths can be followed: in the first one, the perturbed velocity model becomes the starting velocity field for the velocity inversion algorithm; the second path is a short way that involves only migration velocity analysis.

The post-stack split-step Fourier migration method was developed by Stoffa et al. (1990), and was partially implemented in the University of Edinburgh by Thomas Rühl, as a part of the TRACS program. I continued his work at the university, and the computing program was fully implemented on a Sun Sparc station 10 machine in May 1995. After that, I parallelised and implemented the

computer code on a Cray T3D computer at the Edinburgh Parallel Computing Centre (EPCC). The Fourier finite-difference migration by Ristow and Rühl (1994) was incorporated in a further step. The following lines summarise how the program is structured and how it works.

A.2 Theoretical aspects

The split-step Fourier migration method is based on a modification to the phase-shift migration. This method uses a constant reference velocity to migrate the seismic data in the frequency-wavenumber domain. The lateral velocity variations are taken into account as a perturbation term that acts in the frequency-space domain.

Consider the wave equation in a 3D acoustic constant-density medium:

$$\nabla^2 p - u_s^2 \frac{\partial^2}{\partial t^2} p = 0, \tag{A.1}$$

where $p = p(x, y, z, t)$ is the pressure field. The slowness u_s, defined as

$$u_s = 2/c(x, y, z), \tag{A.2}$$

is separated into two terms:

$$u_s(x, y, z) = u_{s_0}(z) + \Delta u_s(x, y, z). \tag{A.3}$$

where $u_{s_0}(z)$ is the reference slowness and $\Delta u_s(x, y, z)$ contains the lateral variations in the slowness.

Transforming equation (A.1) into the frequency domain results in

$$\nabla^2 P + \omega^2 u_s^2 P = 0, \tag{A.4}$$

with $P(r, z, \omega)$ the Fourier transform of $p(x, y, z, t)$. Here r is the horizontal position vector.

After including the slowness variation, the wave equation (equation A.4) becomes

$$\nabla^2 P + \omega^2 u_{s_0}^2 P = -S_f(r, z, \omega), \tag{A.5}$$

with

$$S_f = \omega^2[2u_{s_0}\Delta u_s(r, z) + \Delta u_s^2(r, z)]P(r, z, \omega), \tag{A.6}$$

acting like a source term.

The inhomogeneous wave equation (A.5) is solved in two steps after neglecting the Δu_s^2 term:

1. Considering the upgoing wavefield $P_0(r, z_n, \omega)$ and its Fourier transform $\bar{P}_0$ in the k_r space, a phase shift is applied using the reference slowness u_{s_0}

$$\bar{P}_1(k_r, z_n, \Delta z, \omega) = \bar{P}_0(k_r, z_n, \omega)e^{ik_{z_0}\Delta z}, \tag{A.7}$$

where the reference vertical wavenumber k_{z_0} is

$$k_{z_0} = \sqrt{\omega^2 u_{s_0}^2 - k_r^2}, \tag{A.8}$$

and u_{s_o} is the mean slowness for the interval between z_n and z_{n+1}.

The pressure field $\bar{P}_1(k_r, z_n, \Delta z, \omega)$ is inverse transformed back to the space domain, resulting

$$P_1(r, z_n, \Delta z, \omega) = (\frac{1}{2\pi})^2 \int_{-\infty}^{\infty} \bar{P}_1(k_r, z_n, \Delta z, \omega)e^{-ik_r \cdot r} dk_r. \tag{A.9}$$

2. In this step, a second phase shift that takes into account the perturbation

term in the slowness is introduced (see Stoffa et al., 1990, for more details)

$$P_0(r, z_{n+1}, \omega) = e^{i\omega \Delta u_s(r,z)\Delta z} P_1(r, z_n, \Delta z, \omega). \tag{A.10}$$

The migrated data at z_{n+1} are given by,

$$p(r, z_{n+1}, 0) = (\frac{1}{2\pi})^2 \int_{\omega_1}^{\omega_2} P_0(r, z_{n+1}, \omega) d\omega. \tag{A.11}$$

Ristow and Rühl (1994) point out that the split-step Fourier migration gives accurate results only when smooth variations of the velocity field are used, and when the angles of propagation are approximately vertical. So, they propose a hybrid method, the Fourier finite-difference migration, that combines the split-step Fourier migration with an optimised finite-difference operator that handles strong lateral velocity variations.

They rewrite equation (A.4) for a 2D earth model as

$$\frac{\partial P}{\partial z} = i\sqrt{\frac{\omega^2}{c^2} + \frac{\partial^2}{\partial x^2}} P \tag{A.12}$$

with $P = P(x, z, \omega)$. A second operator that uses the reference velocity c_0 is introduced, and the difference between the full and the approximate operators is expanded to give:

$$\begin{aligned} d_o &= \sqrt{\frac{\omega^2}{c^2} + \frac{\partial^2}{\partial x^2}} - \sqrt{\frac{\omega^2}{c_0^2} + \frac{\partial^2}{\partial x^2}} \\ &\approx \left(\frac{\omega}{c} - \frac{\omega}{c_0}\right) + \frac{\omega}{c}\left(1 - \frac{c_0}{c}\right) \left[\frac{\frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}}{a_1 + b_1 \frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}} + \ldots\ldots \right]. \end{aligned} \tag{A.13}$$

Here a_1 and $b_1$are fitting parameters (Ristow and Rühl, 1994) and $c_0 \leq c(x, z)$.

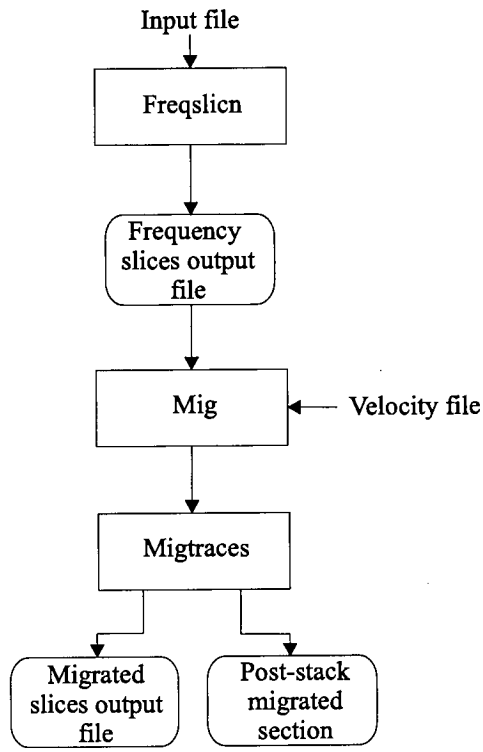


Figure A.1: Flow diagram for split-step Fourier and Fourier finite-difference migrations.

From equation (A.13) the full operator can be written as

$$\sqrt{\frac{\omega^2}{c^2}+\frac{\partial^2}{\partial x^2}} \approx$$
$$\sqrt{\frac{\omega^2}{c_0^2}+\frac{\partial^2}{\partial x^2}}+\left(\frac{\omega}{c}-\frac{\omega}{c_0}\right)+\frac{\omega}{c}\left(1-\frac{c_0}{c}\right)\left[\frac{\frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}}{a_1+b_1\frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}}+\cdots\cdots\right]. \quad \text{(A.14)}$$

The first two terms, corresponding to split-step Fourier migration are the phase-shift operator and the first-order correction term, respectively. The third term represents the finite-difference operator. The main advantage of this method is the fact that the maximum dip migration angle is not fixed (it depends on the velocity values).

A.3 Overview of the computing program

The program has three main steps as shown in Figure A.1. The "Freqslicn" step prepares the input data for further migration. Several parameters are computed. The synthetic traces are built up for a spike operator, otherwise the seismic traces are input as a datafile. The seismic data are Fourier transformed from the time domain to the frequency domain. The number of migration frequencies is calculated as well, and a frequency-slices file is output. The parallel code uses the Message Passing Interface, and this preparation step is performed by the processor *root* as shown in Figure A.2.

In the main step, "Mig", the processor *root* reads the velocity values, and each individual processor reads a frequency slice. The velocities are extrapolated to provide a velocity field for the null traces. There are two main loops: the first one controls the depth steps, and the second one controls the migration frequencies. The shift vectors are prepared considering the reference constant velocity c_{min}, and the data are Fourier transformed from the space domain to the wavenumber domain. The phase shift term is applied, and the inverse Fourier transform sends back the shifted data to the space domain.

A new set of shift vectors is created using the actual velocities and the reference constant velocity. Then, the split-step correction term is applied as a second phase shift. If doing Fourier finite-difference migration, the finite-difference term is calculated and applied as well. In the final step, "Migtraces", the processor *root* sums the calculated wavefields over the frequencies in order to get the wavefield at $t = 0$. The migrated seismic traces and the migrated depth slices are written to separate datafiles.

A.4 Results

In order to check the implementation of the program, a spike at 508 ms at the centre of a 2D model is migrated. The sample rate in time is 4 ms, and 128 traces with an offset of 25 m are considered. The downward continuation step size is 12

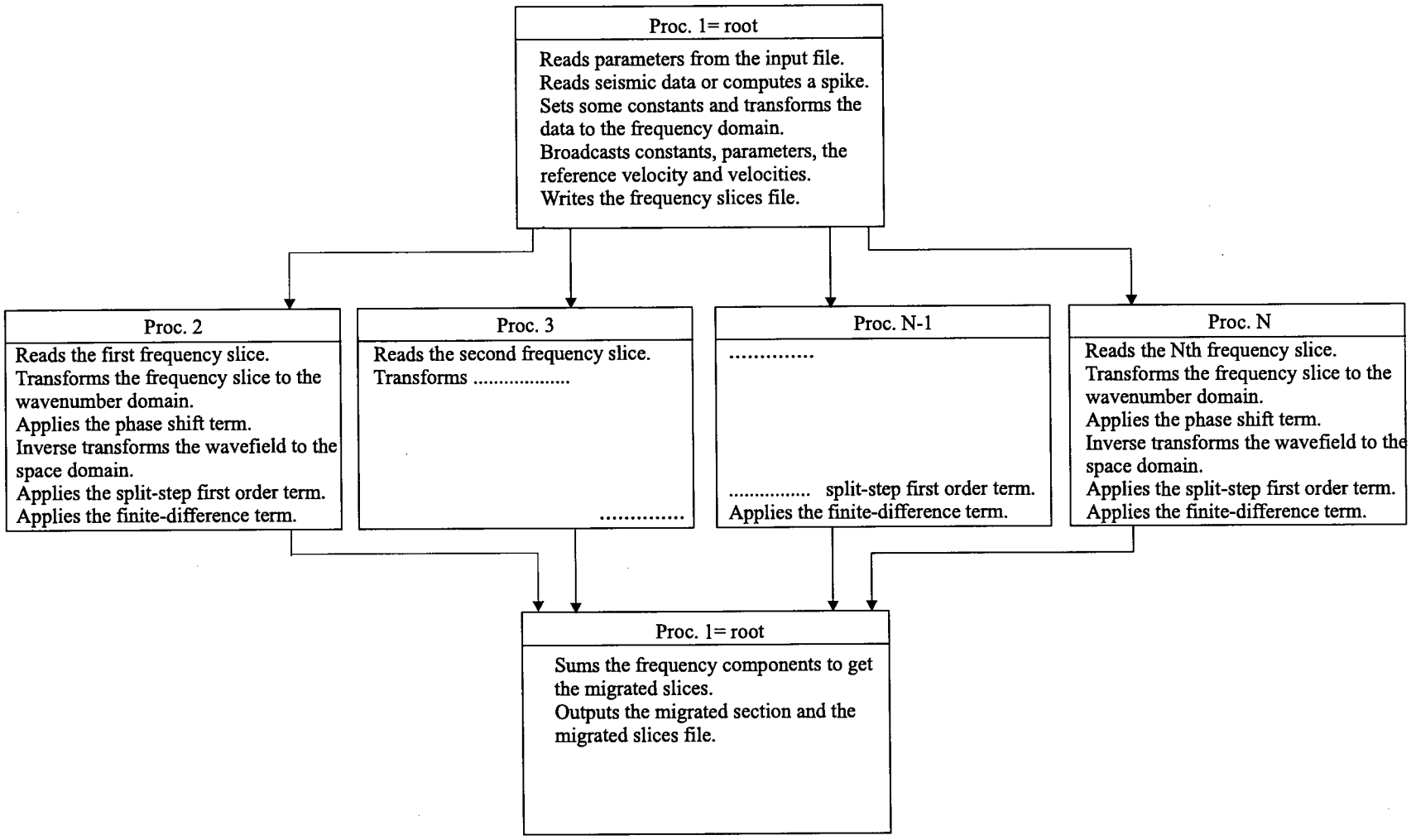


Figure A.2: Computing parallel scheme for split-step Fourier and Fourier finite-difference migrations.

m, and 100 depth slices are calculated. The minimum and maximum frequencies are 1 Hz and 40 Hz, respectively. The velocity values are constant in depth, and increase linearly from the left hand side of the model (2000 m/s) to the right boundary (4000 m/s). Figure A.3 displays the results after applying the split-step operator and the Fourier finite-difference operator, respectively. The execution time in the Cray T3D is at least 10 times faster than the time required by the Sun Sparc workstation.

A.5 Summary

The split-step Fourier migration and the Fourier finite-difference migration are powerful tools to be considered when migrating post-stack seismic data in structurally complex areas. Both migration methods take into account arbitrary lateral velocity variations, and the resolution of the dipping events depends on the varying component of the velocity field. The parallel migration code, implemented on a Cray T3D computer using Message Passing Interface, is at least 10 times faster than the standard code in a Sun Sparc workstation.

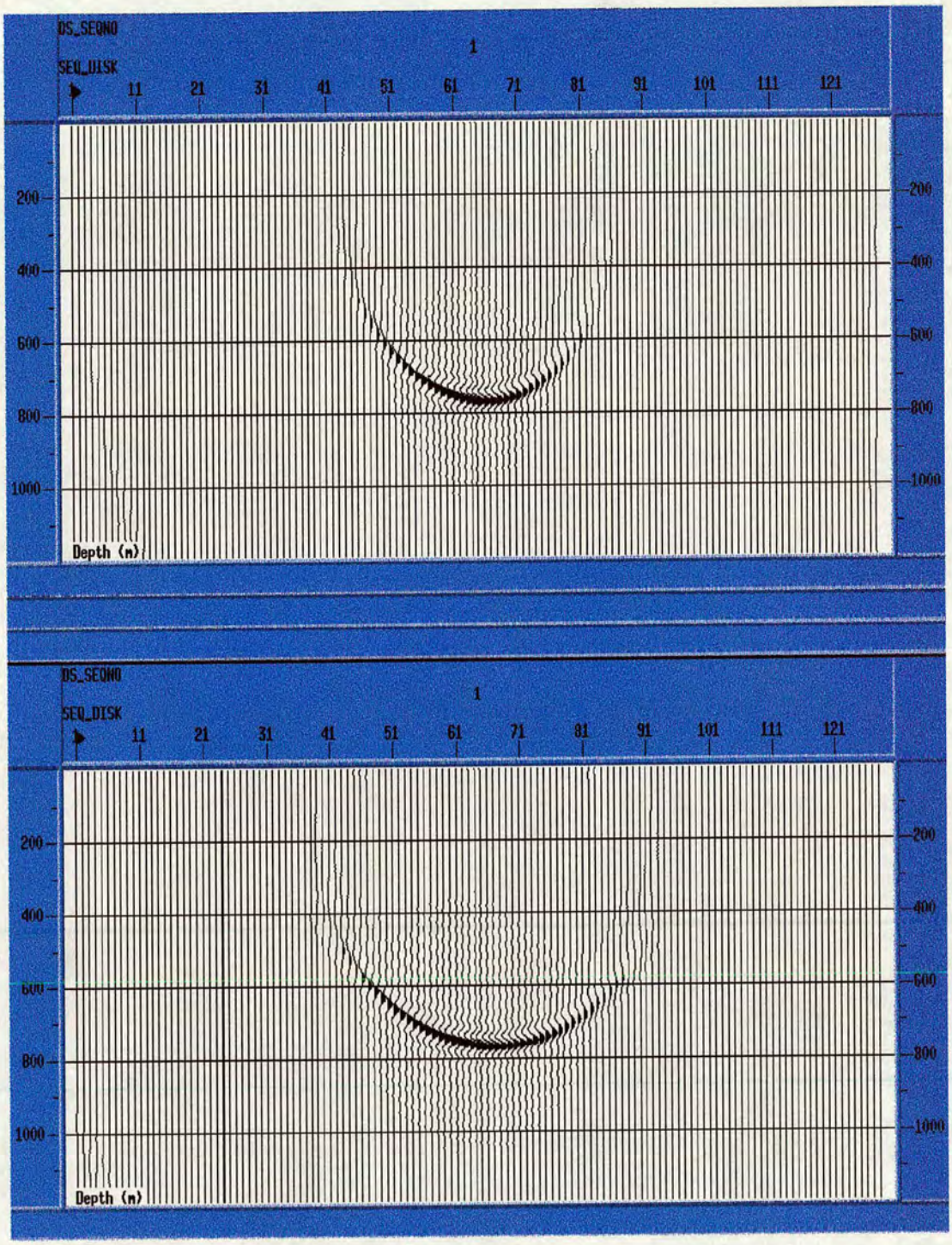


Figure A.3: Migration of a spike at 508 ms. Top: split-step Fourier migration. Bottom: Fourier finite-difference migration.